TABLE OF CONTENTS

No. 1

AN ANALYSIS OF CERTAIN VARIABLES IN A DEVELOPMENTAL STUDY OF LANGUAGE 3
By Florene M. Young

INFANT DEVELOPMENT UNDER CONDITIONS OF RESTRICTED PRACTICE AND OF MINIMUM SOCIAL STIMULATION 143
By Wayne Dennis

AN ANALYSIS OF THE MENTAL FACTORS OF VARIOUS AGE GROUPS FROM NINE TO SIXTY 191
By Benjamin Balinsky

No. 2

FACTORS INFLUENCING PERFORMANCE ON GROUP AND INDIVIDUAL TESTS OF INTELLIGENCE: I. RATE OF WORK . . . 237
By Mary Woods Bennett

INDIVIDUAL DIFFERENCES IN APPERCEPTIVE REACTION: A STUDY OF THE RESPONSE OF PRESCHOOL CHILDREN TO PICTURES 319
By Elisabeth W. Amen

The Journal of Psychology

EDITED BY

Carl Murchison

The Journal of Psychology was founded at the request of many psychologists throughout the world for the purpose of establishing a medium in which publications shall be immediate and in which other publication experiments may be carried out.

(1). The subscription price will be $7 per volume, two volumes per year.

(2). The format of each volume will be such that each subscriber may tear down the assembled volume into its component articles.

(3). Except through the fault of the author, the lag in publication should not be more than three or four weeks.

(4). The author will receive 200 reprints of his paper, and the subscriber will receive the accumulated publications bound into half volume units. The second half of each volume will contain the title page and table of contents for that volume. Each half volume unit will be mailed to the subscribers at its stated date.

(5). Each author will be charged the actual manufacturing costs of any zinc or half tone cuts.

(6). The publication charges will be as follows:

4	page signature	$18
8	" "	24
12	" "	36
16	" "	48
20	" "	60
24	" "	72
28	" "	84
32	" "	96
36	" "	108
40	" "	120

These charges are tentative and will be adjusted as justified by the subscription list (especially by the effect of the war in China and in Europe).

(7). The accumulated units will be mailed to subscribers quarterly: January, April, July, and October. The units may vary in size, to accommodate publication needs.

Two volumes annually. Per annum $14.00; per volume $7.00; per half volume unit $4.00. Complete sets at $7.00 per volume, plus transportation.

THE JOURNAL PRESS
2 Commercial Street
Provincetown, Massachusetts
U. S. A.

$7.00 per volume
Single numbers $4.00

QUARTERLY
Two volumes per year

February, 1941
Volume 23, First Half

GENETIC PSYCHOLOGY MONOGRAPHS

Child Behavior, Animal Behavior, and Comparative Psychology

If this space should be unstamped, this is the regular library edition. But if this space is stamped with a designating title, this is a special edition, sold under the restrictions of a bilateral contract, and may not be resold for a period of five years from date of publication.

FEBRUARY, 1941

(Manuscripts are printed in the order of final acceptance)

AN ANALYSIS OF CERTAIN VARIABLES IN A DEVELOPMENTAL STUDY OF LANGUAGE 3
By Florene M. Young

INFANT DEVELOPMENT UNDER CONDITIONS OF RESTRICTED PRACTICE AND OF MINIMUM SOCIAL STIMULATION 143
By Wayne Dennis

AN ANALYSIS OF THE MENTAL FACTORS OF VARIOUS AGE GROUPS FROM NINE TO SIXTY 191
By Benjamin Balinsky

Copyright, 1941, by The Journal Press
Provincetown, Massachusetts
Entered as second-class matter December 1, 1925, at the post-office at Worcester, Mass., under the act of March 3, 1879
Reentered as second-class matter April 15, 1937, at the post-office at Provincetown, Mass., under Act of March 3, 1879.

Published as a separate and in *Genetic Psychology Monographs*, 1941, 23, 3-141.

AN ANALYSIS OF CERTAIN VARIABLES IN A DEVELOPMENTAL STUDY OF LANGUAGE*[1]

FLORENE M. YOUNG

Jesup Psychological Laboratory, George Peabody College

Acknowledgments 5
I. Introduction and review of related studies 7
II. Procedure and data 13
III. Length of sentence 33
IV. Number of words spoken 55
V. Parts of speech 79
VI. General summary and conclusions 123
Appendix *A* 129
Appendix *B* 131
References 137

---

*Received in the Editorial Office on April 4, 1940, and published at Provincetown, Massachusetts. Copyright by The Journal Press.

[1] Part of a dissertation presented to the Faculty of the Graduate School of George Peabody College in partial fulfillment of the requirements for the degree of Doctor of Philosophy, 1938.

ACKNOWLEDGMENTS

To Dr. Paul L. Boynton the writer owes a debt of gratitude for constructive suggestions, stimulating criticisms, and careful guidance during the progress of this research.

To Dr. S. C. Garrison, whose course on the *Psychology of Language* furnished the initial stimulus for the present study, the writer wishes to express especial appreciation.

To Dr. A. S. Edwards, whose coöperation and assistance in regard to schedules and programs made possible the completion of this investigation, the writer is under great obligation.

To Miss May Zeigler, Miss Margaret McPhaul, and Miss Grace Barnard for their generous aid in connection with numerous duties, and to Miss Helen Bryant who unselfishly gave of her time to aid in checking the accuracy of classifications, the writer is deeply grateful.

FLORENE M. YOUNG

Department of Psychology
The University of Georgia
Athens, Georgia

I. INTRODUCTION AND REVIEW OF RELATED STUDIES

A. INTRODUCTION

An examination of studies of verbal behavior of young children reveals results that differ widely. The variations are outstanding in reports concerning the proportions of the several parts of speech in the conversation. That differences in home environment and in intelligence may account for some of the deviations is a probability suggested by previous experimenters. That differences in the immediate setting may explain a portion of the discrepancies is a possibility which merits study. That numerous divergencies in findings may be due to the fact that investigators based their classifications upon grammars which differed markedly is a point deserving of more emphasis than it has received. Inasmuch as the present investigation involves subjects from two relatively distinct socio-economic groups, groups differing widely in intelligence, and the language records were secured in four different settings, information is available concerning these factors as possible causes of differences in verbal behavior. Grammatical analysis is based upon the works of the noted authority, Dr. Otto Jespersen (41, 42, 43, 44, 45), particularly upon his *Essentials of English Grammar* (45), and comparisons with six previous investigations are made to indicate similarities and differences in findings.

Much emphasis upon adequacy and size of sample is to be found in psychological literature, and this partly explains the fact that the present study is based upon six hours of language observations for each of 74 subjects, a total of 444 hours. Inasmuch as this is not a normative study but an investigation which proposes to study the influence of certain variables, it is seemingly more important to have an intensive rather than an extensive investigation as far as the number of subjects is concerned.

Numerous valuable investigations in the field of language development have been recorded in America and in Europe, yet additional research is needed in this field. McCarthy (59), in enumerating problems awaiting further study, mentioned among others the optimum length and occasion of conversation samples, and the interrelationships between language development and certain environmental influences. Jones and Burks (48, p. 49) likewise indicated that "the influence of such complicating factors as intelligence level, socio-economic status, and the like, have been recognized but not adequately investigated."

It is not strange that psychologists have shown an increasing interest in language, for as Garrison and Garrison (29, p. 202) have well stated, "the problem of language both in its origin and in its development in the individual is essentially a psychological one." In studies of the adjustment of the organism to the environment, language has been recognized as vital both in the history of the race and the individual in making known the needs of the self and the group. The relation of language to the affective life has been emphasized by authorities who, perhaps, in investigating the phylogenetic and ontogenetic aspects of the subject have been impressed by the struggles and vicissitudes arising in the maintenance of equilibrium or the preservation of existence itself. These writers have stated that language had its birth in crises which brought forth "instinctive ejaculations" in primitive man and in conditions which elicited emotional expressions in infants. The emphasis, therefore, has been upon some relationship to the action of the sympathetic division of the autonomic nervous system.

Other investigators, more impressed by the harmonious phases of development in the child and in the race, have emphasized the overflow of energy, the physiological values, the imitation of self or others, and the amusement of oneself or the audience resulting from exercise of the speech mechanism. The emphasis, therefore, has been placed upon some relationship to the action of the parasympathetic division of the autonomic nervous system, or upon the cortical inhibition of the lower centers and the functioning of the cerebral cortex.

Whether language in its phylogenetic and ontogenetic origins had its inception in the activities of the sympathetic division or the parasympathetic division of the autonomic nervous system, or the cerebral cortex, has not been determined; but that the daily employment of language is affected by and related to the total, unified functioning of the integrated organism is a matter of more general agreement. Many of the speech disorders in childhood listed by Seth and Guthrie (77) are illustrative of the deviations which are caused from or accompanied by emotional difficulties. These authors (77, p. 5) emphasized the fact that inefficiency or inadequacy in speech implies "as a necessary consequence an impaired adjustment." Brooks and Shaffer (12, p. 371) stated: "The child's social development is profoundly influenced by his language development." Low (55, p. 7) testified to the "wealth of disintegration phenomena observed in the field of psychotic speech and thought." Studies of language development are important to furnish information concerning any factors which may significantly influence verbal responses and, at the same time, the degree of adjustment of the organism.

Although language may have had an humble origin, consisting perhaps of crying and babbling, in its later manifestations it represents man's loftiest intellectual achievement. That language is necessary to thinking has been disputed, but that the record of the intellectual development of the race and the individual has been recorded largely in terms of language growth and the use of symbols must be admitted. The significance of language as a revelation of maturation and learning, and as a possible cause of mental growth, can scarcely be over-estimated. The far-reaching influence of this instrument is obvious when it is realized that, as Lewis (53, p. 8) indicated, "it subserves all the purposes of human activity—practical, scientific, aesthetic, and religious alike."

B. REVIEW OF RELATED STUDIES IN THE GENERAL FIELD

1. *Socio-Economic Status*

Inasmuch as the Relief subjects in the present investigation came from homes of lower socio-economic status than did the Regular subjects, related studies should be cited to show what differences have been found to exist between such groups. A pertinent description was given in the *Report of the Committee on the Infant and Preschool Child* of which Dr. J. E. Anderson (1) was chairman. Field workers obtained data on 3000 American families, and the report contains charts showing the superiority of the environment of children from the upper economic levels on practically every item, including home ownership, number of occupants per room, health and education of parents, number of toys, books, amount of medical care, and the like. The statement was made that:

> The evidence for the existence of differential environments as far as the child is concerned is overwhelming. When the population is divided on the basis of socio-economic status

> and the practices of the resulting groups are studied in detail, a picture is obtained of a society composed of a series of cultures, *each of which is fairly distinct from but overlaps the one immediately below and above it and all of which seem to be arranged in a linear series.*

The committee concluded that as the child developed practically every aspect of his life was affected and "only a few phenomena seem to be independent of the socio-economic factor."

2. *Effect of Unemployment*

In addition to recognizing the environmental factors which have just been described, it should be recalled that half of the subjects in the present study came from homes in which the depression was having direct and perhaps serious results. In some instances the father was unemployed, and in others he was doing part-time or temporary work. The problem therefore arises as to the possible effects of the depression upon the group which may be termed Relief subjects. Three studies relevant to this question will be mentioned. Jones (47) referred to studies by Buseman and Harder, Buseman and Bahr, and Holzhauer, who investigated the effects of paternal unemployment upon school children. Girls were found to show injurious effects to a more marked degree than boys, and younger children more than older children. Buseman inferred that the effects resulted primarily from poorer hygienic conditions and their influence upon child health and resistance to fatigue. Palmer (72) investigated height and weight and observed that it was the children from families whose incomes had fallen from a higher income to a low level who had been most affected in weight by the depression.

Rundquist and Sletto (76, p. 366), in a study of 3000 cases, found that men who receive relief were not characterized by feelings of inferiority or unfavorable attitudes toward the family. They found, however, that poorer morale was characteristic of the unemployed as a group. "Discouragement and a sense of hopelessness were greatest among men in the semi-skilled and unskilled groups, among older men, and among the men with the least education," according to these authors. The Relief subjects in the present study came from homes in which the paternal type of occupation was semi-skilled, unskilled, or day-laborer in classification; therefore, it may be assumed that the findings which have just been quoted would apply, at least to some degree, in the present instance.

Since the physical, mental, emotional, and social phases of growth appear to be interrelated, it is possible that the effects of paternal discouragement and maladjustment, combined with poorer hygienic conditions, may have influenced the adjustment of the Relief subjects in the present investigation and thereby affected the verbal behavior.

3. *Socio-Economic Status and Intelligence*

The *Twenty-seventh Yearbook of the National Society for the Study of Education* (66) gave an annotated bibliography of studies on socio-economic status and intelligence. More recent summaries relating to occupational level and intelligence have been given in Boynton's (9) and Pintner's (73) able discussions of this subject. Hildreth (38), although noting wide overlapping in mental ability between the several occupational levels, found essentially what previous investigators had indicated, that the *IQ*'s for the professional class were the highest and that intelligence decreased with decrease in the socio-economic level.

In addition to his conclusion that "in general the median scores increased *pari passu* with economic level," Jordan (49) noted that the effect of the poor environment apparently was cumulative, the *IQ*'s being smaller in the succeeding years.

McCarthy (57, pp. 26-31) selected preschool subjects so that the socio-economic classes were represented in their proper proportions. The mean *IQ*'s as obtained for each group ranged from 100.3 for Group VI (Unskilled labor) to 118.1 for Group I (Professional). Goodenough's findings, as stated in the same discussion, were indicative of a similar trend.

Coffey and Wellman (13, p. 193) found that significant differences in intelligence at the time of entrance to preschool were shown to be associated with occupational class even though the two lowest socio-economic groups were not involved in the study.

To emphasize the fact that the relationship between socio-economic status and intelligence is far from perfect, Bayley's (5) study may be mentioned. In an investigation involving 60 children from the ages of one month to 72 months, the correlations were found to be larger after 24 months but they remained below .30. A somewhat closer relationship was noted by Maller (60) with 100,098 school children. He reported a coefficient of .50 when the average *IQ*'s of neighborhoods were correlated with their socio-economic status.

Another angle of the problem has been investigated by Wellman (87, 88, 89). Coffey and Wellman (13) found that during six months of preschool attendance all groups gained five to eight points in *IQ*. On the other hand, Kawin and Hoefer (51, p. 51) observed no significant differences in growth as revealed by mental tests in the fall and spring of the school year, when a nursery school and a non-nursery school group were compared. Kawin and Hoefer (51, pp. 2-3) cited several similar studies the results of which were conflicting.

The subject of language development has been carefully reviewed by McCarthy (59, p. 329) and reference may be made to this exceptionally thorough summary for further material in the general field. Those investigations, or aspects of investigations, which are related specifically to one definite phase of the present study will be discussed later in connection with the chapters on those particular topics.

II. PROCEDURE AND DATA

A. PURPOSE

The present study proposes to investigate the following questions: What variations in language apparently are concomitant with changes in the environmental setting of a group of preschool subjects? Do fluctuations in the amount of verbal behavior, the length of sentences, and the frequencies of the several parts of speech seemingly accompany changes in setting? Do the changes occur in the same manner and degree in two relatively distinct socio-economic groups? How do the subjects in the two groups compare in the various aspects of language development? In what ways do the findings, which are based almost entirely upon the classification found in Jespersen's *Essentials of English Grammar* (45), differ from those obtained by other studies of the parts of speech?

B. SUBJECTS

The subjects are 74 children who were enrolled in the Nursery School of The University of Georgia during one or more quarters of the years 1934, 1935, 1936, and 1937. The basis of selection was merely the fact of attendance in 1934 or 1935 when under-privileged cases and children from families on government relief predominated. During this period the Nursery School was a Federal Emergency Relief Administration project, conducted with the coöperation of The University of Georgia. A few children from homes of the upper economic level attended, the fees being paid by the parents of the subjects. When the Nursery School resumed its original policy of having subjects only from families which paid tuition, the children came from homes of the upper socio-economic levels. The two types of cases are referred to as Relief subjects and Regular subjects. The latter were selected to match the former group, the pairing being primarily on the basis of sex and age. This procedure yielded four groups of subjects: Regular boys, matched for age and sex with Relief boys; and Regular girls, matched for age and sex with Relief girls. From Table 1 it will be seen that there are 20 cases for each of the two groups of boys, and 17 cases for each of the two groups of girls. The parents of the subjects were native-born Americans, and English was the language used in the homes.

1. *Age*

The subjects range in age from 30 to 65 months. Several studies have used younger subjects, but in this instance it was proposed to study cases in whom comprehensible language was assumed to be predominant. McCarthy (57, p. 51) found that 89 per cent of the responses at 30 months were comprehensible, hence no subjects less than that age were selected for the present investigation. Table 1 gives the distribution of subjects. Unfortunately, it was impossible to have the same number of children in each age group. The group 54 months of age was the largest in size, with 18 cases. The matching of groups was such as to make possible comparisons between Regular and Relief boys, Regular and Relief girls, and between Regular and Relief subjects. Although the comparisons may not be quite as accurate, similarities and differences occurring in connection with other group comparisons may be studied, as follows: Regular boys and Regular girls, Relief boys and Relief girls, Regular boys and Relief girls, and Relief boys and Regular girls.

TABLE 1

DISTRIBUTION OF CASES BY AGE, GROUP AND SEX

	Number of						
Age	Boys			Girls			
range	Reg.	Rel.	*All*	Reg.	Rel.	*All*	*All*
30-35	4	4	8	3	3	6	14
36-41	3	3	6	2	2	4	10
42-47	3	3	6	3	3	6	12
48-53	3	3	6	3	3	6	12
54-59	5	5	10	4	4	8	18
60-65	2	2	4	2	2	4	8
Total	20	20	40	17	17	34	74

Table 2 gives the per cent of subjects at each age and the mean for each age group. The means of the columns are based upon individuals rather than upon groups, since the latter are composed of different numbers of subjects. It will be observed that the four groups are very much alike in age. Relief boys are approximately a month older than Regular boys, with whom they are paired, and the largest difference at any age level occurs at 60 months. As will be seen from Table 2, girls are slightly older than boys. However, the means of the age groups 42, 48, 54, and 60 months are higher for

TABLE 2

MEAN AGE FOR EACH GROUP OF SUBJECTS AND PER CENT OF CASES IN EACH AGE GROUP

Group				Mean age of				Per cent in group	
		Boys			Girls				
age	Reg.	Rel.	All	Reg.	Rel.	All	All	Boys	Girls
30	32.6	32.7	32.7	34.5	34.3	34.4	33.4	20	17.7
36	36.8	37.8	37.3	38.4	38.1	38.3	37.7	15	11.8
42	44.1	44.4	44.3	43.4	43.7	43.6	43.9	15	17.7
48	51.4	51.4	51.4	50.6	50.1	50.3	50.9	15	17.7
54	56.9	58.1	57.5	57.1	57.5	57.3	57.4	25	23.5
60	62.7	64.8	63.8	62.4	63.1	62.8	63.3	10	11.8
All	46.9	47.6	47.2	48.0	48.0	48.0	47.6	100	100.2

boys than for girls. All ages have been computed on the basis of the mid-point of the observation period for each child.

2. *Intelligence*

Intelligence was measured by using the *Merrill-Palmer Scale of Mental Tests* by Stutsman (82). This scale includes 13 items of a verbal type, but it is composed mainly of performance tests. Since allowance is made for omissions and refusals, it is especially appropriate for young children. Statements by other investigators concerning this scale may be found in the studies by Driscoll (23, p. 82), Mowrer (65), Kawin (51, p. 16), Bridges (11), and Blackfan (6).

Results obtained from administering the *Merrill-Palmer Scale of Mental Tests* were available for all subjects except two Relief boys. About 40 per cent of the children were tested during the quarter when language records were being made for them. In nearly 20 per cent of the cases it was necessary to use test results which were secured the preceding or the succeeding year. To offset this disadvantage is the fact that scores are more readily comparable when all subjects have been given the same test, and the fact that the same examiner administered all tests.

Percentile ranks are given in Table 3. There is a marked difference between Regular and Relief subjects as far as test results are concerned, the percentile ranks being 70 and 46, respectively. Some of the groups show much variability from age to age. For instance, the mean percentile ranks for Regular boys range from 55 to 82.

TABLE 3

Mean Percentile Rank Obtained on *The Merrill-Palmer Scale of Mental Tests*, by Age, Group, and Sex*

Age range	Boys Reg.	Boys Rel.	Boys All	Girls Reg.	Girls Rel.	Girls All	All
30-35	82	57	70	73	38	56	63
36-41	77	35	56	85	45	65	61
42-47	70	43	57	67	47	57	57
48-53	58	56	57	64	52	58	58
54-59	70	46	58	74	52	63	61
60-65	55	38	47	60	35	48	48
All	69	46	58	71	45	58	58

*Two Relief boys had taken no tests, therefore, the percentile ranks are given for only 18 Relief boys.

3. *Socio-Economic Status*

In Appendix *A* is given the distribution of subjects according to paternal occupation. Inasmuch as most of the Relief subjects came from homes in which the father was unemployed or doing temporary or part-time work, an effort was made to find out what vocations had been engaged in for several previous years. Some information was obtained for all except eight per cent of the cases, and in these instances such facts as the number of years in school and location of residence were used to supplement the information concerning the socio-economic status of the family. The occupational classification was obtained from Goodenough and Anderson (34, p. 237), who state that the classificatory scheme "is in part based upon the Barr and Taussig Scales." The explanation of the classification accompanies Appendix *A*. All of the Regular subjects are to be found in Group I, Professional; Group II, Semi-professional and managerial; or Group III, Clerical, skilled trades and retail business. Classification IV, Farmers, was not used. Although several families owned farms, they did not live on them and the fathers gave most of their time to other occupations. All of the Relief cases are in Groups V, VI, or VII.

Approximately 38 per cent of the Regular subjects in the present study are in Groups I and II, while 62 per cent are in Group III. The mean percentile ranks by paternal occupation are given in Appendix *A*. Occupational Groups I, II, and III have means of 80, 75, and 65, respectively. Occupational Groups V, VI, and VII, in which are found all Relief subjects, have means of 63, 41, and 38, respectively. These data are consistent with those quoted earlier from the results of other investigators, in that Occupational Groups I, II, and III surpass V, VI, and VII in intelligence, as measured by the test specified.

C. Method of Securing Data

1. *Description of Settings*

The records were obtained in the Nursery School in the four types of environment which are arbitrarily designated as Outdoor, Indoor, Dinner, and Picture settings. These terms do not indicate all the implications involved, and when they are used they connote the facts given in the following paragraphs.

When records were secured in the Outdoor setting, the subject was engaged in free play outdoors. He chose his own companions, toys, equipment, and activities without adult interference, unless injury or serious injustice occurred or were imminent. Especial care was taken to prevent adult intrusion when language records were being secured for a subject.

The grounds were spacious, with grass, trees, and shrubbery to make them more attractive. The equipment included the usual playground furnishings, as, slides, swings, parallel bars, ladders, sandpile and sand toys, and boxes of several sizes. In addition, wagons, tricycles, garden tools, and sometimes indoor toys, were brought to the playground. The freedom in choice of materials, activities, and companions, a minimum of interference by adults, and the large areas for free play, were considered as possible factors which might influence the language of the subjects. By the time the subjects went to the playground most of them had had time to become fairly well adjusted for the day, having overcome any temporary difficulties experienced at parting from their parents. In some cases the children seemed to be unusually quiescent, perhaps sleepy, for a time after arrival, but by the time they went outdoors the "warming up" period usually was past.

On the other hand, the Indoor records were sometimes influenced by the factors mentioned, the temporary upset over leaving parents, the fact that favorite playmates had not yet arrived, and the necessity for "warming up" and becoming adjusted. Adult supervision and interference seemed to be more necessary in the Indoor than Outdoor setting. The records included verbal responses made during such routine activities as removal of wraps, combing hair, preparing for weighing and measuring, and the like. Several records for each child were made in a playroom in which there were the usual kinds of indoor equipment, including many types of blocks, dolls, trucks, wagons, and the like. To prevent crowding in the main playroom some of the children usually were playing simultaneously in another room, a factor which reduced the child's choice of playmates. It seems probable that Regular and Relief subjects differed in the time of getting up each morning. From the information available, it appears that Relief subjects must have arisen at least an hour earlier, and this fact may have exerted some influence upon the records of these subjects. The Indoor setting was selected with the thought that the time of day, the location in a more restricted area, more adult interference and contacts, less choice in the number and identity of playmates, and somewhat different types of toys, equipment and routine, might influence the amount and quality of language.

The Dinner setting was even more restricted than the two just described, since the child sat at a table with an adult and from one to four other children. When the number of observers or children warranted it, the subjects were seated at tables in two, sometimes three, different rooms. The subjects were placed so as to reduce problems and unnecessary friction, since one of the main concerns of the staff was to promote desirable reactions in connection with mealtime. The factors just described, combined with the child's relative preoccupation with eating, produced a setting inherently different from the other situations used in the present investigation.

The Picture setting included examination of many different types of picture books, making simple picture puzzles, playing with colored wooden cubes, and putting pegs into several different types of peg-boards. Usually there were from two to five children present in the room and an assistant was directed to show pictures to the child if he did not talk spontaneously. If the child preferred the other activities (puzzles, blocks, or peg-boards) after looking at pictures for a few minutes, the assistant furnished him with other playthings. The assistants also were directed to answer the child's inquiries, encourage talking on his part and even to ask him questions if necessary.

Since the writer was always present to make the records, she gave the cues for most of the questioning. The responses in this setting were, in part, deliberately elicited by adults. The factors which would tend to make this setting somewhat different from the others are as follows: closer contact with adults, the use of questions to produce talking when it did not occur spontaneously, the stimulation of pictures and a few restricted types of play materials which might tend to stimulate enumeration rather than spontaneous language, and the presence of a relatively small number of playmates.

Of the four types of setting the Picture setting may be the most readily compared with McCarthy's (57, pp. 32-34), since she used picture books and toys and, if necessary, elicited responses. The method for the present study was not as standardized as was McCarthy's, inasmuch as the presence of playmates affected the subjects in the present study and in her procedure only the mother and experimenter, or the experimenter alone, were in the room with the child.

Although the language records in the present study were obtained over a period of four years, the physical environment furnished by the Nursery School changed surprisingly little. The same daily program was followed during the four years. The toys and equipment were essentially the same, such replacements or repairs being made that there were no important changes in the nature of the equipment. The same head teacher, Miss Grace Barnard, was in charge during the entire four years, and Miss Catherine Newton, an expert in the field of nutrition, supervised the meals for that period. However, there were changes in the personnel of assistants and observers and in the enrollment from year to year, and even from quarter to quarter.

2. *Number and Duration of Observations*

The experimenter made verbatim records of the meaningful and meaningless vocalizations of each subject. A system of abbreviations and symbols for names and common words was used to facilitate recording. Table 4 indicates the distribution of records by settings. For each case 12 records of 10 minutes in length were made in the Outdoor setting. There were eight, four, and four periods, respectively, for the Indoor, Dinner, and Picture settings. The total number of these uniform periods was 28, and the total time was 280 min-

TABLE 4

NUMBER OF 10-MINUTE PERIODS AND TOTAL TIME OF OBSERVATION FOR EACH OF 74 SUBJECTS IN THE OUTDOOR, INDOOR, DINNER, AND PICTURE SETTINGS

Records	Settings	Periods	Minutes
(*A*) Uniform for all subjects	Outdoor	12	120
	Indoor	8	80
	Dinner	4	40
	Picture	4	40
(*B*) Non-uniform	Miscellaneous	8	80
	Total (*A*)	28	280
	(*B*)	8	80
	All	36	360

utes per child. In addition, miscellaneous records were made (Table 4). For some subjects these were based upon mixed settings. In a few cases all miscellaneous records were made in the Outdoor setting, and in others all were made Indoors or at Dinner. In most instances, those made in the Outdoor situation were the most numerous. Because the miscellaneous settings varied from subject to subject as described, the computation of results has been based largely upon the uniform records where the number of each type of setting was the same for each case. The tables in the succeeding chapters indicate whether they are based upon the uniform records totalling 280 minutes per subject, that is, 345 hours and 20 minutes for the 74 subjects, or upon 360 minutes per child, a total of 444 hours for the 74 subjects.

Discussions in the sections on *Number of Words Spoken* and *Length of Sentence* are based upon uniform records. The material in the section on *Parts of Speech* likewise has as its basis these records, the only exception being the data in and pertaining to Table 21.

As previously stated, the number of uniform periods per subject totalled 28, 12 having been obtained Outdoors, eight Indoors, four at Dinner, and four in the Picture setting. For more accurate comparison of the uniform records in certain portions of the succeeding chapters, the material was made comparable by reducing the number of Outdoor and Indoor periods to four, which was the number of 10-minute observations actually obtained in the Dinner and Picture settings. The mean noun frequencies, for instance, were secured for four Outdoor periods by obtaining the averages for Periods 1, 5, and 9, for Periods 3, 7, and 11, for Periods 2, 6, and 10, and for Periods 4, 8, and 12. By following this procedure, 80 records were available for each of the groups of 20 boys, and 68 records were available for each of the groups of 17 girls, in each of the four different settings. *This procedure was followed only for the uniform records* (Table 4).

The intervals over which the records were made varied in individual cases because of absences, schedules, week-ends, and weather. The shortest time needed to secure the 360 minutes for a subject was 11 days, and the average was 21 days. The mean time was 15 days, the minimum eight days, and the maximum 31 days from the securing of the child's first record, in the case of the uniform records. An especial effort was made to get the records of the subjects in the groups of 30 and 36 months in as short an interval as possible, and none exceeds 19 days.

In all cases of uniform records, a recording in each setting was made during the first three days of observation for a specific child, and in every instance a record of each type was made during the last five days of observation. This precaution tends to offset the possible effects of the duration of the experiment upon the settings.

The present study is based only upon the comprehensible utterances and a report of the non-comprehensible vocalization will be given in another paper.

3. *Advantages of the Procedure*

Fisher (27, p. 10) has indicated the values of securing children's language in a nursery school situation, pointing out that the child is then "in a position where he must adjust to individuals of diverse personality, both adults and children" and that "in competition with his peers each child must become articulate in order to make his wants and desires known."

In testing any skill or set of facts, it is advisable to "sink shafts at various points" and obtain as many samples as possible in representative situations. The present study is based upon data chosen from four different types of setting. The 280 minutes of uniform records, or the total of 360 minutes per subject if the miscellaneous records are included, are considered an adequate sampling of a child's language. Since the uniform records were obtained over an interval of 15 (mean) days, the effects of temporary periods of noncoöperation, fatigue, interrupted schedules, sleep and meals, would tend to be counterbalanced.

Boyd (8) stated that:

> What is needed is a sample of the child's speech sufficiently varied to include all the typical forms of expression. Such a sample can best be obtained by a verbatim record of series of *remarks made when the child is at play, at meals, on a railway journey*, out walking, on visits to friends, at the story hour, engaged in soliloquy, or, indeed, at any period except prolonged spells of emotional disturbances.

The importance of securing records of language in several different settings is understandable in the light of Boyd's statement that "the child uses different words and frames different kinds of sentences with every change of environment." The present study meets the major requirements stated by Boyd.

4. *Sources of Error*

In making observations, words and syllables were sometimes lost because other sounds, particularly the voices of the other children, temporarily obscured the sound of the subject's voice. In the Outdoor setting some difficulty was experienced in making records because of the amount of vigorous activity engaged in at times by the subjects.

It was originally intended that all records should be secured by stenographers, since it was thought that the speech would be most accurately recorded by this procedure. However, comparisons of records made by stenographers with those secured simultaneously by the experimenter revealed that the stenographers who were available for such work failed to understand much of the speech, or relatively large portions were unrecorded since they were unwilling to follow the active subjects from place to place during the observation periods. All records, therefore, were made by the experimenter who used a system of abbreviations and symbols to facilitate recording, and had numerous practice periods before the experimental records were begun. Whenever the experimenter realized that an appreciable portion of an utterance had not been secured, that part of the record was discarded and the situation was treated as an intermission. Although every effort was made to secure accuracy of recording, it should be stated that the method was not ideal. Use of an electric recording apparatus would have been arranged except that the subjects did not stay in one place for long enough periods of time, and this method was not feasible for obtaining the spontaneous language of subjects who usually were playing with other children.

It was impossible to obtain the records at a set interval after the child's entrance into the Nursery School. Inasmuch as some cases appear to be well adjusted after two or three days and others seem to reach that stage only after several weeks of attendance, it would seem useless to set some arbitrary time as an "orientation period" before beginning language records. Even had this been done, the effect of absences still would have to be considered, for the subjects had to readjust, especially after prolonged absences. In the present study, no records were begun before the child had been in attendance for at least two days. Since the recording of language for uniform records extended over a period of approximately 15 days, the possible effect of recent entrances would tend to be diminished. A scrutiny of the four groups fails to show any marked differences between them in the "length of time in school" before the beginning of the language records. In all groups there were cases where the children had been attending for less than a week, and in all there were instances in which the subjects had been in Nursery School for one or two previous quarters. In another section are given the results of certain investigations in which the effects of attendance have been studied.

D. Procedure in Analyzing Data

1. *Authority*

As stated previously, Jespersen's *Essentials of English Grammar* (45) has been used as the basis of classification in the present investigation. Since individuals who are trained primarily in psychology are not necessarily familiar with the names of experts in other fields a brief introduction may be appropriate. Dr. Otto Jespersen has attained international recognition as an authority in the field of grammar. In the introduction to Webster's *New International Dictionary,* Unabridged (1935), is the statement, "Important sources of terms and information have been Jespersen's *Philosophy of Grammar* and *Modern English Grammar.*"

In the preface to *Parts of Speech and Accidence* by Dr. George O. Curme (15, p. ix), is the statement that the author "desires to reacknowledge here his heavy indebtedness to the great *Oxford Dictionary* and the large English Grammars of Jespersen. . . . "

Jespersen's *Essentials of English Grammar* (45) is the volume which has been most used in the present investigation although certain references have been made to the volumes of *A Modern English Grammar* (42, 43, 44), to *The Philosophy of Grammar* (41), and to *Language, Its Nature, Development and Origin* (40), all by Jespersen.

The importance, in fact, the necessity of stating one's authority when making a study which involves grammar will be realized from the following quotation from Kantor (50, p. 5):

> "That grammar is full of difficulties it is not necessary to argue. How many kinds of sentences are there? Some grammarians say three, while others insist upon four. Nor are the three or four the same in the estimation of different writers. And what, by the way, is a sentence? And how do sentences differ from words?" He continues (50, p. 6): "What indeed are parts of speech? Although the ordinary answer is that they are the elements or units of which speech is constituted, the grammarian is unable to stipulate the nature of these units. Unfortunately he lacks a single, adequate criterion to distinguish between one speech part and another. When is a noun a noun and when is it a verb or an adjective?"

In the preface of the *Report of the Joint Committee on Grammatical Nomenclature* (46, p. vi) is the statement: "Nowhere else, it is true, has so great a variation of terminology come into existence as in the grammar of our mother tongue."

To cite an example of the disagreements in this field Jespersen's criticism (41, p. 84) of the *Report of the Joint Committee on Grammatical Nomenclature* may be quoted. He says:

> The term *pronoun* is sometimes restricted (generally in French books, but also in the Report of the Joint Committee on Terminology) to those words which function as what in Ch. VII I shall call "primary words" while *my* is called a "possessive adjective" and *this* in *this book* a "demonstrative adjective." There is, however, not the slightest reason for thus tearing asunder *my* and *mine*, or, even worse, *his* in "his cap

> was new" and "his was a new cap" or *this* in "this book is old" and "this is an old book" and assigning the same form to two different "parts of speech," especially as it then becomes necessary to establish the same sub-classes of adjectives (possessive, demonstrative) as are found in pronouns.

An experimenter following the procedure recommended by the Joint Committee would tend to have a larger number of adjectives than would be true if Jespersen's views were adopted. Apparent differences in the findings concerning the relative frequencies of the various parts of speech may be due in part to following the statements of different grammarians and not to significant differences in the language used by the subjects. To illustrate this fact, a tabulation including 12 hours of observation from the records of 12 subjects was made for the following words: *my—mine, our—ours, your—yours, his, her—hers, their—theirs,* according to Jespersen's viewpoint and then according to the statement of the Joint Committee (46, p. 15). According to the latter many of these words are to be regarded as adjectives, as described in Jespersen's (41, p. 84) criticism which has just been given. The application of the procedure recommended by the Joint Committee caused the majority of the words to be changed to this part of speech. *My* and *your* are adjectives in this classification and these words have by far the greatest incidence of any which were selected for the comparison. Because pronouns such as *I, me,* and *you* have high frequencies the fact just emphasized may be somewhat obscured in a study of the parts of speech.

Inasmuch as Jespersen's classification of pronouns, other than personal pronouns, likewise should be illustrated, a few examples of sentences taken from the present study will be given, with references to Jespersen's *Essentials of English Grammar* (45) for the specific types listed.

1. "*The same* is the ordinary pronoun of identity" (45, p. 171). *I want the same box you had.*

2. "*Other* may be called the pronoun of difference" (45, p. 179). *Do you want another color?*

3. "The pronoun of unspecified quantity" (*some*) (45, p. 180). *He's ready for some "zeit" (dessert).*

4. "The pronoun of indifference" (any) (45, p. 181). I never had any glasses.

5. "Pronouns of totality" (all) (45, p. 184). I'm gona (going to) stay all day.

Jespersen has pointed out some of the inconsistencies in grammar, as may be illustrated from *The Philosphy of Grammar* (41, p. 88):

> If now we turn to such words as *on* or *in*, we find what is to my mind an exact parallel to the instances just mentioned in their employment in combinations like "put your cap on" and "put your cap on your head," "he was in" and "he was in the house"; yet *on* and *in* in the former sentences are termed adverbs, and the latter prepositions, and these are reckoned as two different parts of speech. Would it not be more natural to include them in one class and to say that *on* and *in* are sometimes complete in themselves and sometimes followed by a complement (or object)?

Although recognizing and calling attention to these inconsistencies, Jespersen (45, p. 69) in his later book accords with common practice. He gives examples: "Is John in? (adv.)"; "in the house (prep.)."

These quotations have been given to illustrate the fact that in the present investigation, the later book is followed when there are divergent views presented in other writings. Exceptions to this practice are mentioned.

To give the reader an idea of a basic viewpoint as stated by Jespersen a quotation from *Essentials of English Grammar* (45, p. 71) will be given:

> This shows that in order to find out what class a word belongs to it is not enough to consider its form in itself; what is decisive is the way in which the word in connected speech "behaves" towards other words, and in which other words behave towards it. If we find that one and the same form is used now as a substantive, now as an adjective or verb, this does not mean that the distinction between word-classes is obliterated in English, for in each particular combination the form concerned belongs decidedly to one class only; but the form should not be looked at in isolation.

Another quotation, taken from the preface to *Essentials of English Grammar* (45, p. 5), will show that the viewpoint held by the authority in question is particularly appropriate in connection with ontogenetic studies of language. He stated that: "It has been my endeavor in this work to represent English Grammar not as a set of stiff dogmatic precepts, . . . but as something living and developing under continual fluctuations and undulations, something that is founded on the past and prepares the way for the future, something that is not always consistent or perfect, but progressing and perfectible—in one word, human."

2. *Specific Methods*

In ascertaining the number of words spoken in a given unit of time, the number of words per response, and the frequencies of the various parts of speech for the several groups and settings, arbitrary rules were formulated in some instances. In most cases, however, the viewpoint of Jespersen was ascertained from his writings and adopted as the procedure. On a few points it seemed advisable to modify the procedure which would have been consistent with his statements. These cases are mentioned in the succeeding enumeration of rules which were followed. The experimenter accepts full responsibility for the arbitrary nature of the statements which are not based upon the views of an accepted authority.

1. Only comprehensible words are treated in connection with the number of words spoken per unit of time, and per sentence, and the frequencies of the several parts of speech. Meaningless syllables are tabulated in separate tables.
2. Contractions of subject and predicate and verb and negative, such as *you're* and *won't,* were scored as two words and both portions classified as parts of speech.
3. Each part of the verb phrase was treated as a separate word.
4. Hyphenated words were scored as single words. The decision as to whether a word should be hyphenated was based upon Webster's *New International Dictionary,* Unabridged (1935).
5. The definition of a sentence as given by Jespersen (45, p. 106), was used as the basis of procedure. He states that "A sentence is a (relatively) complete and independent unit of communication (or—in case of soliloquy—what might be a communication were there someone to listen to it)—the completeness and independence being shown by its standing alone or its capability of standing alone, *i.e.,* of being uttered by itself." He includes amorphous sentences (45, pp. 105-106) and says "it is best to acknowledge amorphous sentences on the same footing as other sentences."

In the present study "sentence" and "response" are used as synonymous terms. When "the length of response" is discussed, therefore, the "length of sentence" is referred to.

6. So-called compound nouns such as "James Russell," "stone wall," and the like, were regarded as two words. Jespersen (45, p. 85) states that: "Substantives are used as secondaries in the innumerable loose collocations in which each substantive preserves its stress and is therefore felt as a separate unit." The reader is also referred to (45, pp. 19 and 31) for other statements relative to the subject.

7. Preschool subjects often use childish substitutes for ordinary words. These were classified as nouns if used as nouns or as other parts of speech if so used. Jespersen illustrates this tendency (40, p. 108) and enumerates such substitutes as "taters" for potatoes, "chine" for machine, "tina" for concertina, and "tash" for moustache.

8. Ellipsis is common in the language of children. In the present study only the words which were actually spoken were counted and classified.

9. Even though some of the words were superfluous, as *"I'm ain't goin',"* they were counted and classified.

10. In cases of metanalysis, the proper division was made according to adult standards. Of this phenomenon Jespersen says (42, pp. 141-142) "Each child has to find out for himself in hearing the connected speech of other people, where one word ends, and the next one begins, or what belongs to the kernel and what to the ending of a word, etc." "Familiar instances are *a naddre* which through metanalysis becomes *an adder, North Thriding,* which becomes *North Riding.*" Further discussion will be found in *Language, Its Nature, Development and Origin* by Jespersen (40, p. 173).

11. In cases of the infinitive with *to*, the *to* was classed separately as a preposition on the basis of Jespersen's (45, p. 345) statement that "the preposition *to* no more belongs to the infinitive as a necessary part of it, than the definite article belongs to the substantive. . . ."

12. Although the term *substantive* is preferred by Jespersen (45, p. 66), the word *noun* is used in the present study for the sake of convenience.

13. Pronominal adverbs are classified with adverbs, although they partake of the characteristics of both parts of speech in Jespersen's (45, p. 68) classification.

14. In some of the tables a combination of infinitives, participles, and gerunds is presented. In other tables an arbitrary grouping of adjectives, infinitives, participles, interjections, and the gerundial construction, is given. These combinations were due to the necessity for reducing the number of classifications for certain parts of the investigation. In all cases the tables indicate clearly the headings which have been combined.

15. In certain tables, in order to show possible development in each, the adverbs, prepositions, and conjunctions were tabulated separately. In other instances they were combined under the heading *particles,* used by Jespersen (45, pp. 68-69). The procedure which was followed will be found in connection with each table.

16. Jespersen (45, p. 174) classifies articles with pronouns. In the present study, in order to show possible development in their use, they have been tabulated separately in some tables. In all cases the statement is made in connection with the table if articles are included with pronouns.

E. ACCURACY OF SCORING

In classifying the parts of speech for the mass of data which formed the basis of the present investigation precautions were exercised to insure accuracy. After careful study of *Essentials of English Grammar* by Jesperson (45) the experimenter classified all the parts of speech in all the records. Preliminary attempts had shown the necessity for formulating the list of specific rules which have been enumerated. In addition, a supplement was made for the index to *Essentials of English Grammar* by Jespersen (45), with illustrations of various classifications taken from the records of the subjects. This facilitated the scoring process and enabled the experimenter to be much more consistent than otherwise would have been possible.

A check upon the accuracy of scoring was the comparison of the experimenter's marking of the same records, a total of 54 hours of observation, in September and again in December. Numerous other records had been scored in the meantime. The December classification was made without reference to the previous work. The disagreements were due largely to change of opinion as to the use of hyphens, the reversal of decisions as to whether an utterance should be classified as comprehensible or incomprehensible, and to change of viewpoint concerning the grammatical classification of certain words. The re-scoring of these records, which were the first that had been scored, corrected the errors where they were the most numerous. The writer made all the frequency distributions of the parts of speech, and in tabulating the records which had not been re-scored as described, relatively few errors were found. Each part of speech was examined for classification by the experimenter on at least two different occasions. In addition, the debatable words were scrutinized many times over the period of more than two years when most of the scoring was done.

F. Explanation of Terms

Throughout the investigation the word "setting" is used to refer to one of the four settings: Outdoor, Indoor, Dinner, or Picture. The term "group" refers to one of the socio-economic groups: Regular boys, Relief boys, Regular girls, or Relief girls. When the expression "age group" occurs, reference is being made to one of the age sub-divisions of one of the socio-economic groups as, the subjects 36 months of age, or all the subjects of a given age. As previously stated, "uniform records" are those in which the settings have equal weight, and for which the total time per subject is 280 minutes (Table 4). The "non-uniform" records also include miscellaneous records, most of which were made in the Outdoor setting. When these are included the observation time totals 360 minutes for each subject. The tables specifically indicate the fact if these data constitute part of the tabulations.

Certain abbreviations have been used, as follows: *Out.*—Outdoor; *In.*—Indoor; *Din.*—Dinner; *Pict.*—Picture; *Reg.*—Regular; *Rel.*—Relief; *M.*—Males; *F.*—Females.

All references to ages in the tables are given in terms of months. The age groups contain subjects scattered throughout the range which is indicated. For instance, the age group, 30-35 months, may have cases in any or all of the six months which are included.

The central tendencies, deviations, and critical ratios may be found in the tables in the main part of the study, while many of the correlations are given in Appendix *B*. In the self-comparisons of the several groups of subjects in the four settings, the formula used was that for the reliability of the difference between two means that are correlated, as recommended by Cooke (14, p. 106). The 168 product-moment correlations used in connection with this formula are given in Appendix *B*. Statistical analysis has been based mainly upon formulae taken either from the text by Cooke (14), or from that by Garrett (28).

III. LENGTH OF SENTENCE

A. RELATED STUDIES

1. *Reliability of Sentence Length*

Investigators in the field of child psychology have long regarded the sentence length as one of the most valuable indices of linguistic development, but measurement of the reliability of this criterion is a comparatively recent emphasis.

McCarthy (57, p. 46) correlated the odd- with the even-numbered responses in order to see how consistently the children used sentences of a certain length, and obtained a correlation of .91.

Smith (78, p. 17) stated concerning length of response that: "This failure to gain after four years and the extreme variability in the older children of the same mental ability make it seem probable that the sentence length as a measure of sentence development has no significance after four and one-half years." Smith also pointed out the fluctuations due to changes in "the activity of the child, his state of excitement, and the subjects he is addressing."

Nice (67), in recording 13 hours of conversation for her daughter, found the range of mean sentence length from hour to hour to be from 3.5 words to 4.5 words.

Williams (90, p. 13) obtained .86 as the reliability of sentence length in his investigation.

Davis (17) used subjects of ages five and one-half, six and one-half, and nine and one-half. For these ages she obtained reliability coefficients for the longest sentence ranging from .59 to .86; for the mean of five longest sentences the reliability coefficients ranged from .84 to .92; for the number of one-word remarks the range was from .79 to .94. The mean of 25 sentences proved to be the most reliable at each age, with reliability coefficients ranging from .87 to .95.

2. *Sex Differences*

McCarthy (57, p. 54) found that girls were slightly superior in the length of sentence used at most age levels. Her findings are given in Table 11 in comparison with those of the present study. She considered the differences between the sexes as "suggestive and possibly significant although they do not meet the statistical criterion of the significance of a difference." She called attention to the fact that "girls may go through the developmental cycle more rapidly than do the boys, but the boys practically equal them at the close of this rapid developmental period."

Smith (79, p. 209) found in comparing boys and girls that "at two years the girls' sentences were significantly longer than the boys'" and that "except at five years the girls' sentences averaged slightly longer than did the boys' sentences."

Davis (16, p. 49) obtained a superiority in length of sentence for girls as compared with boys at ages five and one-half, six and one-half, and nine and one-half years, indicating that sex differences are not confined to the preschool level in this criterion of language development. Davis (17, p. 72) also compared the long, short, and mean sentences as measures of linguistic growth and concluded that the superiority of girls becomes "even more clear when long sentences are the method of measurement."

Fisher (27, p. 79) found for a highly selected group of 72 children that the mean number of words per response for the girls was greater at each age except at 42 and 54 months, at which levels the boys excelled.

3. *Socio-Economic Differences*

McCarthy (57, p. 57) found that Occupation Group I was clearly superior to all other groups in length of sentence used. For the combined 24 and 30 months age groups the mean length of sentence used by Group I was 5.7 words, but in Groups V and VI the means were only 2.0 and 1.8, respectively. The differences between the upper and lower occupational groups were "statistically significant in all but the 18 months group."

Smith (79, p. 204) compared children of three groups: (*a*) children from the preschool laboratory and similar background, (*b*) children from the free kindergartens and skilled labor groups, (*c*) the day nursery children and those with similar background. She stated that: "Class I in every case and at every age and in both situations used longer sentences than did the other two classes."

Fisher (27, p. 6) had 72 subjects from the Professional and Managerial Groups and a comparison of the mean number of words per response with the data from Smith and McCarthy, both of the latter studies having relatively heterogeneous groups, shows the marked superiority of Fisher's cases over those in the other studies.

Fisher's (27, p. 80) subjects, for instance, had sentences with a mean length of 4.8 words at 24 months, while the subjects in McCarthy's and Smith's studies for that level averaged 1.8 and 1.7 words, respectively.

Stalnaker (80) obtained language records from 14 West Virginia subjects from Occupational Groups I, II, and III. Her findings also indicate the superiority of subjects from such socio-economic classes, inasmuch as her records clearly surpassed those of Smith's subjects at the ages investigated.

Davis (16, p. 61) found for her subjects of five and one-half, six and one-half, and nine and one-half, that children from the upper occupational groups made longer remarks than children from the lower occupational groups. At five and one-half years the girls from the lower occupational groups made shorter remarks than the boys, but at nine and one-half years the boys from the lower occupational groups tended to use the shorter sentences.

Little and Williams (54, p. 76) obtained data on sentence length, with many other measures of language development, for 285 children of different socio-economic levels. They emphasized the fact that: "Below normal mental age accompanied by an orphanage environment appears to lower linguistic achievement at all ages included in this study. No conclusions are drawn regarding how much the retardation is due to lowered mental age and how much to environmental influences."

4. *Age Differences*

Smith (78, p. 17) found a gradual increase in length of sentence for 88 subjects through four and one-half years. In a later study including 305 records, Smith found an increase at each age from 18 to 68 months in the length of response, although the increase after four and one-half was small at each six-month interval. The correlation for age and length of sentence was .29 ± .043, with mental age and questions held constant.

Fisher (27, p. 80), using a selected group from the upper occupational groups, found a lengthening of the mean response from 3.7 words at 18 months to 9.5 at 54 months. Stalnaker (80) likewise obtained a rapid increase in length of remark with age for 14 subjects of two through four years of age. Williams (90, p. 14) secured a correlation of .68 ± .04 between chronological age and length of sentence.

Nice (68) suggested that the length of sentence, as it developed with age and as affected by other variables, was a criterion of speech growth with four main stages: single words; early sentences averaging more than one and less than three words; short sentences of three to four words; a transition stage, and the better established sentence of six to eight words.

Day (19) found an increase in the length of response with age, singletons being superior to twins in length of sentences used. The superiority became more marked with age. McCarthy (57, p. 54) likewise noted a lengthening of the sentence with age, the mean lengths at 18 and 54 months being 1.2 and 4.6 words, respectively. With subjects of five and one-half, six and one-half, and nine and one-half years of age, Davis (16, p. 49) obtained conclusive evidence that the sentence increased in length beyond the preschool age. The means for the subjects of the ages stated were 4.6, 5.3, and 6.5 words per sentence, respectively. Day, McCarthy, and Davis used the same procedure in obtaining data.

McCarthy (58, p. 296) reviewed many investigations and concluded that: "These studies all seem to indicate that at these early ages the mean length of response is a significant measure of the developmental process. There is a marked increase with age which is most rapid up to three and one-half years."

5. *Differences in Intelligence*

Williams (90, p. 14) secured a correlation of .69 ± .04 between mental age and length of sentence. Smith (79, p. 210) reported a smaller correlation, a coefficient of .39 ± .04, between sentence length and mental age, with questions and chronological age held constant. Davis (16, p. 55) obtained correlations between sentence length and intelligence at five and one-half, six and one-half, and nine and one-half years as follows: .48 ± .03, .21 ± .08, and .20 ± .06.

6. *Differences in Setting*

McCarthy (56) used 31 children two years and four months to four years and eleven months of age as subjects when she recorded language in two different situations: with the experimenter, indoors, and with other children, outdoors. In both settings, 50 consecutive remarks were recorded. The mean length of response for the boys indoors was 3.61 and while outdoors it was 3.94 words. The girls showed almost the same result in both environments, 4.41 and 4.45. That the subjects did not actually respond alike, however, is indicated by a correlation of .54 as obtained between the mean lengths of response in the two situations. In conclusion, McCarthy stated: "This study indicates quite clearly that, for work with groups, the situation has little effect upon the mean length of response, but that in order to get an adequate measure for any individual child it is necessary to obtain samplings of conversations in several different situations."

Smith (78, p. 18) emphasized the importance of control of setting, as follows: "This makes it evident that it is necessary, if sentence length is to be considered, that the children be in comparable situations, for example, engaged in reasonably active play with their fellows, with very little adult attention."

Smith (79) investigated the differences between lengths of sentence where the child was at home with adults (Situation *A*) and where he was engaged in play with other children (Situation *C*,) using 152 records of boys and 153 records of girls of 18 to 68 months of age. At each age level, the sentence lengths were greater in Situation *A*. Smith stated that "this difference is statistically significant at four and five years as well as for the whole series." Smith's procedure differed from McCarthy's in that in Smith's study the observer did not address the child unless necessity arose; whereas, in McCarthy's investigation some of the responses were elicited by questions and consequently the percentage of answers, usually shorter than average sentences, was higher in McCarthy's study. Smith (79, p. 189) also pointed out other possible explanations of the differences in results of the two studies. For the 20 children who were observed by Smith (79, p. 190) in both Situations *A* and *C* the mean sentence length was 4.4 and 3.5, respectively. In Situation *A* more compound and complex sentences, more questions, and fewer single-word, negative, and imperative sentences occurred.

Anderson (2) studied 150-word passages written by students whose mean age was approximately 19 years, and concluded that "within the language product of a single individual indices of written language vary with the situations in which language is used and with the subject matter."

7. *Individual Differences*

Due to one or more of the factors which have been mentioned, or to other variables which may be as important, individual differences in sentence length are striking. Davis (17) found for 436 subjects that about 20 per cent of the sentences were one word in length, while there were many sentences that reached 20, 30, or even 50 words. The range in length was from one word to 56 words. The mean of all 50 remarks at age five and one-half was 4.57, but the mean of the five longest remarks was 10.3, and the mean of the longest sentences was 13.5 words.

Nice (70) stated that the average length of sentence used by the Brandenburg's daughter at age three was 6.6 words and at age four, 7.5 words; by Kirkpatrick's daughter at four, 7.0 words; by Boyd's daughter at ages three, four, and five, 6.4, 6.9, and 7.5 words, respectively. These means are much superior to those found by McCarthy (57) and Smith (78).

McCarthy (57, p. 62), Day (19), and Davis (16, pp. 52-62; 18) have demonstrated the superiority of the child who has much contact with adults, and the superiority of only children over twins and singletons. Davis (18) stated that "onliness increases the length of sentence as much as fourteen months of chronological age."

8. *Differences in Other Factors*

Portenier (74) used 25 preschool children in a Federal Emergency Relief Administration Nursery School and found that the "well-adjusted children" surpassed those who were in the "mediumly well-adjusted group" in mean sentence length. These two groups excelled the "poorly adjusted cases" although the latter were older chronologically and mentally. The mean sentence lengths for the three groups, which were 41, 39, and 49 months of age, respectively, were 3.2, 2.8, and 2.6 words.

B. ANALYSIS OF DATA

The data in this chapter are based upon analysis of the uniform records (Table 4) which included 280 minutes per subject. As previously described, four records from each setting were used for each subject. This was made possible by obtaining the means for four Outdoor periods, for instance, by averaging Periods 1, 5, and 9;

Periods 3, 7, and 11; Periods 2, 6, and 10; Periods 4, 8, and 12.

Only comprehensible verbal behavior is treated in the present section since it seems inadvisable to put meaningless syllables on a par with words. The incidence of non-comprehensible vocalization will be given in a later paper.

1. *Reliability of Sentence Length*

Inasmuch as the length of reponse is regarded as an important index of language development, the reliability of the measure is of interest. The Pearson product-moment correlation technique was used to obtain correlations between the odd-even-numbered periods for the Outdoor and Indoor settings. These two situations constitute twelve and eight, respectively, of the 28 uniform periods, as shown in Table 4. The data referred to in Table 5 were divided by the odd-even method, and since this method takes account of only one-half the actual data, the coefficients were corrected by using the Spearman-Brown prophecy formula. The results, shown in Table 5, indicate satisfactory reliability. The lowest correlation, .78±.033, occurs at the 60 months level, and the highest, .93±.010, at the age of 36 months. The mean sentence length may be considered a reasonably reliable measure of language development in the present study.

TABLE 5

RELIABILITY COEFFICIENTS OF LENGTH OF RESPONSE FOR AGE GROUPS (BASED ON OUTDOOR-INDOOR UNIFORM RECORDS)

C. A.	Pearson *r*	*P E*
30	.88	±.014
36	.93	.010
42	.89	.014
48	.88	.016
54	.86	.015
60	.78	.033

2. *Differences Accompanying Changes in Setting*

In Tables 6 and 7 are given the "Central Tendencies and Deviations for Number of Comprehensible Words per Response" as spoken by Regular boys, Relief boys, Regular girls, and Relief girls, in the several settings. The respective means for Regular and Relief boys in all settings are 4.56 and 3.98 words per response. The longest

TABLE 6

CENTRAL TENDENCIES AND DEVIATIONS FOR NUMBER OF COMPREHENSIBLE WORDS PER RESPONSE SPOKEN BY REGULAR BOYS AND RELIEF BOYS IN THE OUTDOOR, INDOOR, DINNER, AND PICTURE SETTINGS

	Regular boys			Relief boys		
Setting	Mean	*SD*	SD_m	Mean	*SD*	SD_m
Out.	4.60	1.34	.15	4.10	1.52	.17
In.	4.60	1.44	.16	4.11	1.28	.14
Din.	4.52	1.32	.15	3.96	1.34	.15
Pict.	4.51	1.29	.14	3.73	1.27	.14

TABLE 7

CENTRAL TENDENCIES AND DEVIATIONS FOR NUMBER OF COMPREHENSIBLE WORDS PER RESPONSE SPOKEN BY REGULAR GIRLS AND RELIEF GIRLS IN THE OUTDOOR, INDOOR, DINNER, AND PICTURE SETTINGS

	Regular girls			*Relief girls*		
Setting	Mean	*SD*	SD_m	Mean	*SD*	SD_m
Out.	4.90	1.45	.18	4.89	1.43	.17
In.	4.89	1.38	.17	4.79	1.30	.16
Din.	4.86	1.21	.15	4.59	1.21	.15
Pict.	4.83	1.05	.13	4.73	1.34	.16

utterances occur in the Outdoor and Indoor settings, and the shortest in the Picture setting. The respective means for Regular and Relief girls in all settings are 4.87 and 4.75 words per response. The longest responses are to be found in the Outdoor and Indoor settings. The shortest for Regular girls occur in the Picture situation, and for Relief girls in the Dinner setting.

In Appendix *B* are given "Correlations between Mean Sentence Lengths Used in the Compared Settings by Regular and Relief Boys." The 12 correlations range from .88 to .95 and indicate a marked tendency for subjects who use long sentences, for instance, in the Outdoor setting, to have similar types of utterances in the other settings.

In Appendix *B* correlations for Regular and Relief girls also are given. They range from .90 to .94, and likewise indicate that sentence length in any one setting tends to be markedly similar to that used by the subjects in the other situations of the present investigation.

In Table 8 are given the "Critical Ratios of the Differences between Specified Settings in the Number of Comprehensible Words per Response for the Group Specified." The correlations involved in the formula may be found in Appendix *B* and the means, standard deviations, and standard errors of the means in Tables 6 and 7. The standard error of the mean was used as the measure of reliability, and a critical ratio of 3.00 may be taken as indicative of a statistically significant difference. In giving comparisons of settings, the situation which has the larger mean will be mentioned first.

In the self-comparisons of Regular boys (Table 8) in the several settings, no differences satisfy the criterion of statistical reliability, the largest critical ratio being 1.92. However, the odds are between 88 and 97 in 100 that there are true differences between the Indoor-Dinner, Outdoor-Dinner, Indoor-Picture, and Outdoor-Picture settings when the specified comparisons are made.

For Relief boys (Table 8), three of the six critical ratios indicate completely reliable differences as a result of the self-comparisons in the several settings, as follows: Indoor-Picture; Outdoor-Picture; and Dinner-Picture. In addition, there are 96 and 99 chances in 100 that there are differences greater than zero which are not due to chance, resulting from the Outdoor-Dinner and Indoor-Dinner comparisons.

TABLE 8

CRITICAL RATIOS OF THE DIFFERENCES BETWEEN SPECIFIED SETTINGS IN NUMBER OF COMPREHENSIBLE WORDS PER RESPONSE FOR THE GROUP SPECIFIED
(*Correlated Observations*)

Setting	In.	Pict.	Din.	Setting	In.	Pict.	Din.
	Regular boys				Relief boys		
Pict.	1.53			Pict.	6.03		
Din.	1.21	0.21		Din.	2.38	3.90	
Out.	0.00	1.92	1.33	Out.	0.13	4.74	1.73
	Regular girls				Relief girls		
Pict.	0.77			Pict.	0.94		
Din.	0.48	0.48		Din.	3.64	2.12	
Out.	0.14	0.89	0.63	Out.	1.35	2.54	4.76

In the self-comparisons of Regular girls in the four settings, no critical ratios indicate reliable differences, the largest being 0.89. The differences between means do not exceed 0.07 words.

The length of verbal response used by Relief girls seems to be more affected by change of setting than that of any other group, except Relief boys. Two of the six critical ratios indicate statistical significance, and they are the result of comparing the Dinner setting *with one of the other settings*, as follows: Outdoor-Dinner; and Indoor-Dinner. The comparison of the Outdoor-Picture situations results in virtual significance, there being approximately 99.5 chances in 100 that a true difference greater than zero exists. The critical ratios for the Outdoor-Indoor, and Picture-Dinner comparisons indicate 91 and 98 chances in 100 that there are true differences which cannot be accounted for by chance.

If the observations are treated as if they are uncorrelated, the critical ratios in Table 8 are reduced in size. Those for Regular boys and Regular girls already are indicative of statistically insignificant differences between the compared settings, and with the uncorrelated observations as a basis the largest critical ratio is 0.44. Relief subjects (Table 8), with correlated observations used as a basis, have five critical ratios indicative of reliable differences between the compared settings. When the observations are treated as if uncorrelated, the largest critical ratio is 1.92.

In Table 9 may be seen the percentage of one-word utterances in the various settings for the several groups. In the Picture situation, the one-word responses comprise 16.9 per cent of the utterances for boys and 12.2 per cent for girls. In the other settings the proportions are less, the respective percentages for Outdoor, Indoor and Dinner records being 8.7, 10.2, and 8.6. It is possible that

TABLE 9

PERCENTAGE OF ONE-WORD RESPONSES BY AGE, GROUP, SEX, AND SETTING

			Boys			Girls		
Setting	*C.A.*	Reg.	Rel.	All	Reg.	Rel.	All	All
	30	18.5	23.6	21.1	17.0	19.9	18.5	19.8
	36	14.0	19.3	16.7	12.3	15.7	14.0	15.4
	42	10.3	16.5	13.4	10.7	9.0	9.9	11.9
	48	7.9	8.4	8.2	6.2	8.6	7.4	7.8
	54	8.2	9.4	8.8	5.9	7.2	6.6	7.7
	60	7.5	7.9	7.7	6.4	8.7	7.6	7.7
	All	11.1	14.2	12.7	9.8	11.5	10.3	11.7
Out.	All	8.6	10.8	9.7	6.4	8.8	7.6	8.7
In.	All	8.2	12.7	10.5	9.9	9.7	9.8	10.2
Din.	All	6.7	9.9	8.3	7.8	9.8	8.8	8.6
Pict.	All	15.4	18.3	16.9	11.2	13.2	12.2	14.6
All	All	9.7	12.9	11.4	8.8	10.4	9.6	10.5

the question-answer incidence in the Picture setting may account for the higher proportions in that situation. For "All" settings combined the proportions for Regular boys, Relief boys, Regular girls, and Relief girls are, respectively: 9.7, 12.9, 8.8, and 10.4.

In Table 10 are given the critical ratios of the differences between settings for "All Subjects." In this case, the method of computing the critical ratios was not the same as that employed in connection with those given in Table 8, inasmuch as the critical ratios for "All Subjects" in Table 10 are not based upon a formula which involves the correlation between settings. Undoubtedly, the critical ratios would have been larger, if the same procedure as was used for the results in Table 8 had been followed. With the method which was employed, no critical ratio exceeds 0.86, and it is apparent, for the combined groups of "All Subjects," that the length of response does not change greatly from setting to setting in the present investigation. However, it should not be forgotten that five of the 24 self-comparisons of specific groups (Table 8) result in statistically significant differences, and nine other such comparisons indicate relatively large differences. It appears that under certain conditions changes in setting may be concurrent with marked fluctuations in the length of response. It should be noted that the critical ratios which satisfy the criterion for statistical significance occur only for Relief subjects, and only when the Dinner setting, in the case of Relief girls, or the Picture setting, in the case of Relief boys, is compared with other settings.

In summarizing, it should be restated that the Dinner and Picture settings may be associated with statistically significant changes in sentence length as compared with other settings, in the case of Relief subjects. Regular boys change markedly in length of response when certain changes are made in setting but the differences do not satisfy the criterion for statistical reliability. Regular girls, who surpass all other groups in sentence length, show very slight changes concurrent with shifts in setting, in this criterion of language development. Apparently the four groups were not identically affected by the four different settings in which they were placed. Judging from the results of the combined groups of "All Subjects" (Table 10), studies of sentence length made in the Picture setting, a situation somewhat similar to that used by McCarthy (57), may be expected to yield findings representative of those which would be secured in any of the other three settings, if the subjects are of both sexes and come from widely differing socio-economic levels, as in the present investigation. If, however, a relatively homogeneous group is being studied it seems advisable to secure responses in several different types of situation. This appears to be especially important if the subjects are of low socio-economic status.

3. *Variations Accompanying Differences in Socio-Economic Status or Sex*

Table 10 shows the "Critical Ratios of the Differences between Compared Groups in Mean Number of Comprehensible Words per Response in the Settings Specified." Comparisons of sentence length used by Regular and Relief boys indicate statistically significant differences only for records made in the Picture setting, but the critical ratios resulting from comparisons in the other three settings indicate from 98 to 99.7 chances in 100 that there are true differences between the two groups. There is not a statistically significant difference between them when they are compared on the basis of the combined settings, although the critical ratio of 2.81 indicates about 99.74 chances in 100 that a true difference exists. In this group comparison, the socio-economic status differs and the sex factor is constant.

When Regular girls and Regular boys are compared, the socio-economic status being relatively constant and sex differing, none of the critical ratios for specific settings indicates the presence of a statistically significant difference. However, the odds are between 89 and 96 in 100 that there are true differences between the two groups, the girls having the longer sentences. For the combined (All) settings, the critical ratio of 1.48 means that there are approximately 93 chances in 100 that the true difference is greater than zero.

Group comparisons show Regular boys and Relief girls to have relatively similar records as regards length of response. The latter excel Regular boys in all comparisons. However, the largest critical ratio is only 1.26, and that for the combined settings is but 0.86. The similarity of the two groups is especially noteworthy since they differ both in sex and in socio-economic status.

The largest critical ratio occurs in the Dinner setting and it indicates approximately 90 chances in 100 that there is a true difference greater than zero when Regular and Relief girls are compared.

No other critical ratios exceed 0.55, which is the result of comparing the two groups in the combined settings.

In spite of the fact that they belong to the same socio-economic level, Relief girls and Relief boys differ markedly in mean sentence length. The former excel Relief boys, and the comparisons of records in specific and combined settings result in statistically significant differences in all instances.

The largest differences between any two groups are those from comparisons of Regular girls and Relief boys. Regular girls greatly surpass Relief boys, and all critical ratios indicate highly significant differences. For the combined settings, the critical ratio is 4.29. The greatest difference occurs in the Picture setting, with a mean sentence length of 4.83 for Regular girls, and 3.73 words for Relief boys. In this group comparison both sex and socio-economic differences are present.

Do the combined groups of girls surpass the two groups of boys in mean length of response? Girls excel boys, but the critical ratio of 2.62 (Table 10) does not satisfy the criterion for statistical significance, although it indicates approximately 99.5 chances in 100 that the true difference is greater than zero. The differences between the two groups of girls and Relief boys are marked, since all comparisons, both in specific and combined settings, result in reliable differences. The differences between the two groups of girls and Regular boys are statistically unreliable, since no comparisons in specific or combined settings result in critical ratios greater than 1.68.

Do the combined groups of Regular subjects surpass Relief cases in mean length of response? Regular subjects are superior, but the critical ratio of 1.76 (Table 10) is not indicative of a statistically significant difference between the two socio-economic levels. The differences between the two groups of Regular subjects and Relief boys are marked, since five of the eight comparisons in specific settings result in reliable differences. The differences between the two Regular groups and Relief girls, however, are negligible, inasmuch as the largest critical ratio resulting from comparisons in specific or combined settings is 1.29.

It is of interest to observe the instances where sex differs and the socio-economic factor is relatively constant. This is true for Regular girls-Regular boys and the critical ratio is indicative of an insignifi-

TABLE 10

CRITICAL RATIOS OF THE DIFFERENCES BETWEEN COMPARED GROUPS IN MEAN NUMBER OF COMPREHENSIBLE WORDS PER RESPONSE IN THE SETTINGS SPECIFIED

Setting	Out.	In.	Din.	Pict.	All	Setting	Out.	In.	Din.	Pict.	All
Regular boys—Relief boys						Regular girls—Relief girls					
Out.	2.17					Out.	0.04				
In.		2.33				In.		0.43			
Din.			2.67			Din.			1.29		
Pict.				3.90		Pict.				0.48	
All					2.81	All					0.55
Regular boys—Regular girls						Relief boys—Relief girls					
Out	1.30					Out.	3.29				
In.		1.26				In.		3.24			
Din.			1.62			Din.			3.00		
Pict.				1.68		Pict.				4.76	
All					1.48	All					3.71
Regular boys—Relief girls						Relief boys—Regular girls					
Out.	1.26					Out.	3.20				
In.		0.83				In.		3.55			
Din.			0.33			Din.			4.29		
Pict.				1.05		Pict.				5.79	
All					0.86	All					4.29
All subjects						Boys—Girls					
Out.	——	0.09	0.64	0.86	——	All					2.62
In.			0.55	0.76	——	Regular—Relief					
Din.				0.19	——	All					1.76

cant difference, on the basis of the combined settings. A similar type of grouping is found for Relief girls-Relief boys, and in this case the critical ratio of 3.71 is sufficient to show statistical significance (Table 10).

It is likewise important to note the instances in which the socio-economic factor varies and the sex factor is constant, as for Regular boys-Relief boys. The comparison made on the basis of the combined settings indicates about 99.74 chances in 100 that there is a true difference greater than zero. A similar type of grouping is found for Regular girls-Relief girls, and none of the differences for specific or combined settings is statistically reliable, the largest critical ratio being 1.29.

It is apparent that effects of sex and socio-economic status do not adequately explain the differences existent in the length of response, since the findings are not consistently indicative of a positive relationship. In this connection it should be again emphasized that the critical ratios for the comparisons, in the combined settings, of Regular girls-Relief boys, and Relief girls-Regular boys, with sex and socio-economic differences existent in both pairs, are 4.29 and 0.86, respectively. Neither sex nor socio-economic status can be the basic explanatory factor when the comparisons of similar types of groupings result in such different findings.

4. *Differences Accompanying Variations in Age*

Previous investigators have found an increase with age in the length of verbal response, and typical findings have been cited earlier in the chapter.

In Table 11 is given the "Mean Length of Verbal Response for the Present Study as Compared with McCarthy's Results." It will be noted that the means are not identical with those obtainable from Tables 6 and 7. In the latter tables, all ages are combined into one group and the results presented for the four settings. In Table 11 the results are given by ages, and since the number of subjects is not exactly the same in the six age groups the means resulting from use of the two procedures are not always the same.

For the individual and combined groups there is a gradual increase with age in sentence length. The mean for the present study at 30 months is 3.2 words, and at 60 months it has increased to 5.2 words. At all ages, Relief subjects are surpassed by Regular subjects and boys are excelled by girls.

McCarthy's results also are given in Table 11 and it will be observed that the means of the present study are somewhat larger than those obtained in her investigation. This may be due in part to the fact that the age groups in the two studies are not identical. McCarthy's (57, p. 25) groups were so selected that each child was observed within a month and a half of the age at which he is classified. In the present investigation the age group, 30-35 months, includes subjects in any months within that range. In addition, the difference in settings may account for some of the differences in the results. This seems probable inasmuch as it has been shown that three of the four groups in the present study have their shortest responses in the Picture setting, and seven of the twelve comparisons of this setting with other situations result in marked or statistically significant differences (Table 7).

TABLE 11

MEAN LENGTH OF VERBAL RESPONSE FOR THE PRESENT STUDY AS COMPARED WITH MCCARTHY'S RESULTS ACCORDING TO AGE, SEX, AND GROUP

Group	Boys			Girls				McCarthy's results		
CA	Reg.	Rel.	All	Reg.	Rel.	All	All	Boys	Girls	All
30	3.3	2.8	3.1	3.4	3.1	3.5	3.2	3.2	3.1	3.1
36	3.6	3.0	3.3	4.1	3.7	3.9	3.6	3.1	3.8	3.4
42	4.4	3.9	4.2	4.5	4.3	4.6	4.4	4.2	4.4	4.3
48	4.9	4.2	4.6	5.1	4.4	4.8	4.7	4.3	4.4	4.4
54	5.0	4.3	4.7	5.4	4.6	5.0	4.9	4.6	4.7	4.6
60	5.2	4.5	4.9	5.9	5.0	5.5	5.2			

Another factor which should not be overlooked is that of the rules for counting words. For instance, "*won't*" was classified as two words in the present investigation, while McCarthy regarded it as one word.

Smith's (79) means for similar groups are 2.5, 3.5, 4.3, 4.6, 4.9, and 5.0 words, per sentence. The subjects were classified so that the group of 30 months includes the range, 27-32 months.

In Table 9 may be seen the percentage of one-word remarks in the various settings for the several groups. The younger subjects have a much higher proportion of this length of response, the decrease with age being found in each of the four groups. For the combined group of subjects at 30 months of age, 19.8 per cent of all remarks are one word in length, but for the 60 months age level only 7.7 per cent are of that length. The greatest decrease comes in the interval between 30 and 42 months, although the four groups do not show the same rate of decrease.

5. *Differences Accompanying Variations in Intelligence*

Differences in length of utterance accompanying variation in intelligence may be observed by comparing the percentile ranks for the several groups for which differences have been obtained in the various comparisons. The percentile ranks were secured from administering the *Merrill-Palmer Scale of Mental Tests* (82) to the subjects, as previously described. The comparisons, based on the critical ratios

TABLE 12

Percentile Ranks Obtained on *The Merrill-Palmer Scale of Mental Tests*, Means for the Groups, and Critical Ratios of the Differences Between Groups for Number of Comprehensible Words Per Response*

Groups compared	Mean length of response		Percentile ranks		Critical ratios
Regular—Relief	4.70	4.33	70	46	1.76
Girls—Boys	4.81	4.26	58	58	2.62
Reg. B.—Rel. B.	4.56	3.97	69	46	2.81
Reg. G.—Reg. B.	4.87	4.56	71	69	1.48
Reg. G.—Rel. G.	4.87	4.75	71	45	0.55
Rel. G.—Rel. B.	4.75	3.97	45	46	3.71
Rel. G.—Reg. B.	4.75	4.56	45	69	0.86
Reg. G.—Rel. B.	4.87	3.97	71	46	4.29

*In naming the groups and in stating means and percentile ranks, the results of the superior group are given first.

in "All" settings, are stated with the percentile ranks in Table 12.

Regular subjects, with a mean percentile rank of 70, surpass in mean sentence length Relief cases for whom the percentile rank is 46, and the critical ratio of 1.76 means that there are 96 chances in 100 that the true difference is greater than zero. Regular boys, with a percentile rank of 69, surpass Relief boys whose rank is 46, the critical ratio of 2.81 indicating approximately 99.74 chances in 100 that the difference cannot be accounted for by chance. Also, Regular girls, with a percentile rank of 71, excel Relief boys whose percentile rank is 46, the critical ratio being 4.29. However, similar comparisons between Regular girls-Relief girls, with respective percentile ranks of 71 and 45, result in statistically insignificant differences. In addition, the fact should be noted that Relief girls-Relief boys, with practically the same percentile ranking, differ significantly in mean sentence length.

Apparently, intelligence as measured in this study is not a consistent determiner of the mean length of utterance. The variables are numerous and it is difficult, if not impossible, to separate them sufficiently to warrant definite conclusions as to causes. Earlier in the chapter were given the relationships between length of response and mental age, or between length of response and intelligence, as obtained by four investigators. The reported correlations ranged from .20 to .69, indicating wide differences in the findings.

6. *Differences Accompanying Variation in the Number of Siblings*

Writers have disagreed as to the possible effect of the number of siblings upon the language development of preschool subjects, therefore it is a matter of interest to observe any possible relationships in the present study.

It should be stated that the criterion of an only child in this investigation is a record of no siblings less than 15 years of age at the time of the language recording. The incidence of only children in the four groups is as follows: Regular boys, 55 per cent; Regular girls, 53 per cent; Relief boys, 10 per cent; and Relief girls, six per cent.

Regular subjects far surpass Relief cases in the incidence of "onliness," but the critical ratio from comparison of length of response fails to satisfy the criterion for statistical reliability. It indicates, however, 96 chances in 100 of the existence of a true difference. Again the close resemblance of the records of Regular boys and Relief girls composed, respectively, of 55 per cent and six per cent only children, must be noted. Relief girls and Relief boys have very similar records of "onliness" in their respective groups, but there is a statistically significant difference between them in length of response.

To attack the problem from another angle, the number of siblings for each group may be counted. The groups had the following numbers of siblings who were less than 15 years of age at the time of the language recording: Regular boys, 12; Regular girls, 9; Relief boys, 31; and Relief girls, 23. Relief subjects far surpass Regular subjects in number of siblings. The possible effect of sibling incidence is indicated by the fact that Regular girls, with a total of nine siblings, significantly surpass Relief boys, who have a total of 31 siblings, in mean length of utterance. Also, Regular boys excel Relief boys in mean sentence length, the critical ratio of 2.81 indicating approximately 99.74 chances in 100 of a true difference greater than zero. However, Relief girls are significantly superior to Relief boys, although the two groups have almost the same average number of siblings. Regular girls and Relief girls, with 9 and 23 siblings, respectively, are more alike than any other groups in mean length of utterance. Almost as great a similarity exists between Regular boys and Relief girls, with 12 and 23 siblings, respectively. From this analysis, it appears that the number of siblings is not a factor which exerts a consistent influence upon mean length of response.

7. *Differences Accompanying Variation in Other Factors*

Responses vary in length from one word to 39 words, the latter sentence occurring in the record of a boy of 53 months. A Relief boy whose maladjustment will be described in a later chapter had the shortest mean sentence length. Although classified in the 42 months age group, his responses averaged only 2.4 words in length, a record inferior to that of the subjects 30 months of age. His maladjustment apparently influenced length of response and also the number of words spoken per 10-minute period.

Factors such as fatigue, unusual home schedules, weather, length of school experience, and state of excitement, probably affect the length of response in some instances. These factors have not been controlled or measured in the several settings. Some of them have been referred to in the description of procedures and settings.

It is assumed that individuals are distributed somewhat according to the normal probability curve when adequate samples of the mean length of response are obtained. The implications of such an assumption have been clearly stated by Edwards (24). The means for the several preschool ages probably fall below those of older subjects. Davis (16, p. 49) indicated that at six and one-half and nine and one-half years the average lengths of response were 5.3 and 6.6, respectively, an increase over the results of the present study.

C. SUMMARY

1. The reliability of the observations of the length of the response is indicated by correlations between odd-even-numbered records (uniform Outdoor and Indoor records) which are .86 ± .015 or higher at all ages except at the 60-months level where the coefficient is .78 ± .033.

2. Correlations between any two of the settings are .88 or higher, indicating a marked degree of consistency in the length of response used by the subjects in the different settings.

3. The mean length of response obtained in the combined settings is as follows for Regular boys, Relief boys, Regular girls, and Relief girls respectively: 4.56, 3.98, 4.87 and 4.75 words.

4. The length of response shows variations which apparently are concomitant with changes in settings, as follows:

(*a*) Of the 24 self-comparisons of specific groups in the several settings only five result in statistically significant differences.

(*b*) From six self-comparisons of "All Subjects" as a combined group no reliable differences result, and the largest critical ratio is 0.86. Apparently the length of response is not greatly influenced by changes in setting.

5. The length of response shows variations which apparently are concomitant with changes in socio-economic status or difference in sex, as follows:

(*a*) Girls surpass boys, and Regular subjects surpass Relief subjects. The critical ratios of 2.62 and 1.76 indicate, respectively, 99.5 and 96 chances in 100 that a true difference exists.

(*b*) Where sex differs and the socio-economic factor is relatively constant, or where the socio-economic factor differs and sex is constant, the results are not consistently indicative of the possible influence of these factors. Comparisons of Relief boys with other groups tend to give statistically significant differences. Comparisons of Relief girls with Regular subjects result in very small differences.

6. The length of response shows variations which apparently are concomitant with age differences, as follows:

(*a*) With increase in age the mean length of response shows an increase for each of the four groups.

(*b*) The proportion of one-word responses shows a marked decrease with age, especially at the early levels.

(*c*) At all ages girls excel boys, and Regular subjects excel Relief subjects. Relief boys have the poorest records at all age levels.

7. The length of response shows variations which apparently are concomitant with differences in intelligence or variations in the number of siblings, as follows:

(*a*) Regular subjects, with a percentile rank of 70, are markedly superior to Relief subjects whose percentile rank is 46. However, marked inconsistencies from other group comparisons indicate that intelligence, as a possible causative factor, does not exert a uniform influence in these instances.

(*b*) Regular subjects, composed of a much larger group of only children and having few siblings as compared with Relief children, markedly surpass the latter in mean length of response. Nevertheless, the contradictions which result from other group comparisons indicate that sibling incidence, as a possible causative factor, does not exert a uniform influence in these instances.

8. Individual records reveal that certain variations in the length of sentence may be attributed to, or may be concurrent with, serious maladjustments.

IV. NUMBER OF WORDS SPOKEN

A. Related Investigations

Interest in the amount of language used by young children has been manifested by investigators who have recorded the number of words spoken during specified periods of time by individual subjects. Waddle (86, p. 174) summarized the well-known studies by Gale, Brandenburg, and Bell. Others will be described in the brief review of certain relevant investigations which follows.

1. *The Time-Sampling Technique*

The accumulation of data for relatively large groups has lagged because of the lack of a reliable method which is more rapid than that usually employed. Goodenough (31) in 1928 described a time-sampling technique developed by Olson. In using this procedure, Goodenough (32) obtained a correlation of .59 from correlating the scores of two observers. She stated that "the low reliability is probably a function of the different hours used for observation. The records of Observer *G* were taken entirely during the lunch hour; those of *M* during the morning free-play hour."

Robinson and Conrad (75) observed 50 cases 24 to 58 months of age for 20 days each, using a form of the time-sampling method, with an entry each quarter-of-a-minute period. The experimenters obtained reliability coefficients of .95 to .99 on split-half correlations for 320 quarter-minute observations per child. They concluded that by "the expenditure of a considerable yet reasonable amount of time, good measures can be obtained by the time-sampling technique for at least the great majority of Nursery School children."

Arrington (3, p. 77) correlated the average number of remarks recorded per five-minute period with the average number of five-second intervals in which talking occurred. The correlation for the 11 children common to both experiments was .84±.06.

Olson and Koetzle (71) tested the time-sampling technique with children who were 46 to 69 months of age, from the upper occupational groups. They used a mechanical hand-tally and a time-out stop watch. The reliability coefficients for the 15 odd and 15 even observations were .93 and .83, when corrected, for the kindergarten and nursery, respectively. The correlations between teachers' ratings of talkativeness and the scores were .40 and .73 for the nursery school and kindergarten, respectively. The mean amount of language spoken by the subjects was 16.5 words per minute. However, the experimenters included "any words, exclamations, or jumbles of sounds produced by the children in word-play." In addition, it should be noted that the subjects had a mean *IQ* of 124.

A similar observation may be made concerning the vocal output of Fisher's subjects (27, p. 80). If the "Average Remarks per Hour" are used in connection with the "Mean Number of Words per Response" it appears that her subjects talked considerably more than those in some investigations. Whether this fact was entirely the result of the selection of the group (Professional and Semi-Professional as to occupation, and mean *IQ* of 136), is not known. Probably some of the difference may be due to the inclusion of "non-verbal or incomprehensible remarks," in the total remarks. Other studies based only upon the number of comprehensible words would necessarily differ to some extent. Fisher (27, p. 79) obtained a correlation of .68±.04 for the total number of remarks on two records taken on different days.

2. *Sex Differences*

Olson and Koetzle (71) found that girls talked more than boys during given periods of time but Fisher (27, p. 68) "found no significant differences between the sexes." Both investigations involved groups which were above the average in intelligence. Goodenough's (32) investigation yielded findings which were similar to those just mentioned, for she obtained no evidence of a sex difference in the group of nursery school subjects for the trait of talkativeness. Davis (16, p. 25) noted that "ratings revealed a consistent tendency toward greater talkativeness in boys than girls" in her study. Her subjects, 436 in number, were older than those in the other three studies and selected so as to be representative of the various occupational groups of the locality.

Inasmuch as it appeared from casual observation in the present investigation that the groups of boys were more active physically than girls and this factor may have affected the vocal output and the accuracy of recording, the findings of Fales (26) are relevant. Using 32 preschool children, 16 boys and 16 girls paired according to chronological age, Fales (26) found that the difference in the vigorousness scores for the play of boys and girls was not statistically significant. The data were reliable as shown by a correlation of .92 between the odd and even five-minute periods. Contrary findings as to sex differences in the vigorousness of play were reported by Goodenough (32) who stated that: "Support for the common belief that boys are more active than girls is found in this study." The difference between the sexes in physical activity was statistically significant.

3. *Socio-Economic Differences*

Gesell and Lord (30) compared 22 preschool children, half from homes in which the fathers were members of the professions and half from homes of poorly favored economic status. They found that "in virtually every field the advantage seems to be with the favored group, whether verbal, practical, or emotional abilities are considered." They further stated that the "most outstanding net disparity between the two groups consisted in the very noticeable difference in the amount of conversation." Of the subjects from the less fortunate group they said that "only twice were comments comparable in spontaneity, if not in content, heard in the *B* Group." This observation is in agreement with that of Markey (62, p. 130) who found that the imaginative play of children of lower socio-economic status was somewhat more "prosaic" than the play of children of higher status, it being more closely connected with their every-day experiences than that of more privileged children.

Davis (16, p. 25) noted for her 436 subjects that the children from the upper occupational groups were more talkative than those of the lower occupational groups. The results were obtained from use of a seven-step scale for scoring of talkativeness. Similar findings were obtained by Goodenough (32) who noted "a progressive decrease in the talkativeness scores as we go down the line of occupational classes."

4. *Age Differences*

Olson and Koetzle (71) found that "within each group there is some tendency for the older children to talk more than the younger." The coefficients between amount and age were .34±.14 and .28±.15 for the nursery school and kindergarten, respectively. Smith (78) found an increase in the amount of language as age increased. Her findings will be discussed more at length in connection with those of the present investigation. Although the subjects used by Fisher (27, p. 68) were not representative of the preschool population it should be noted that she also obtained a "significant positive correlation between the number of remarks per hour and chronological age." The increase with age ceased, however, after the beginning of the fourth year.

5. *Differences in Intelligence*

Fisher (27, p. 68) found "no significant relationship between talking and intelligence, as indicated by a Pearson *r* of .20±.07." It should be recalled, however, that her group was highly selected. The subjects used by Olson and Koetzle (71) also were above the average, with a mean *IQ* of 124. These authors found that intelligence quotients correlated .14±.11 with amount of talking and —.02±.12 with rate. They concluded that: "In the absence of contrary evidence obtained with larger numbers it is tentatively assumed that no judgment of a child's mental capacity at these ages and in the upper ranges of mental ability can be made on the quantity or rate of his talking." A closer relationship was found by Goodenough (32) who obtained correlations between talkativeness and mental age, as follows: .49 for the Minnesota Test and .46 for the Merrill-Palmer.

Mattson's data also contained evidence of a relationship between talkativeness and mental age for the Merrill-Palmer subjects in the investigation. Mattson's study, described by Markey (63, p. 93), obtained for the group whose average mental age was 42 months a mean of 3.21 words spoken per minute, and for the group whose mental age was 66 months, a mean of 9.42 words per minute.

6. *Individual Differences*

Olson and Koetzle (71) found with their subjects that "the child who talks the most will talk approximately seven times as much as the child who talks the least during a given unit of time." Fisher (27, p. 67) noted that the average number of remarks per hour ranged from 23 to 192. She (27, p. 16) stated that the "number of different remarks obtained for the experimental group ranged from 200 for one of the most taciturn two-year-old boys, to 1,728 for a socially aggressive four-year-old boy."

Arrington (3, p. 155) observed that differences were conspicuous in frequency of talking. She (3, p. 154) found, however, that "a high degree of consistency characterized the behavior of the majority of the children whose activities were observed in two successive stages of development."

Individual differences in the amount and rate of talking may be observed in the time required by McCarthy (57, p. 67) to obtain 50 remarks per child. The range was from seven to 50 minutes with a mean of 19.3 minutes. In the study by Davis (16, p. 26) the range was from five to 45 minutes with a mean of 13.6 minutes.

Waddle (86, p. 173) was impressed by the amount of speech used by young children, stating that "children of the preschool period talk most of the time they are awake." On the contrary, Arrington (3, p. 154) found that "talking, or to use the more accurate term, vocalization, occurred on an average in less than one-fifth of the play time at the two-year level, and about one and one-half times as frequently among the three-year-olds."

7. *Differences in Setting*

Ellesor (25) observed a girl 36 months of age (*IQ* 156) for 30 minutes daily for seven weeks in five "free-play situations" and four "constrained situations." She found that the differences between the amounts of vocalization in the five free-play situations were unreliable, and indicated that "the quantity of vocalization had reached a constant level in these situations." However, the critical ratios for the differences between sandpile and home setting, and between sandpile and nursery school with children, were above 2.5. Ellesor stated that the difference in quantity of vocalization between any one of the five "free-play situations" and any one of the four "constrained situations" was found to be reliable. In addition, she obtained "a reliable difference in the quantity of vocalization between any two of the four 'constrained situations.' "

Nice (67) summarized studies by Gale, Bell, Nice, and Brandenburg for a total of 10 children whose ages ranged from two through five years. These children in home settings had records of 10.6 to 21 words spoken per minute. Markey (63, p. 84) called attention to the contrast between these data and the findings of Mattson who recorded for nursery school children a mean vocal output of 6.7 words per minute for subjects ranging in age from 36 to 56 months. Markey stated: "How much difference there is between verbal activity at school and at home it is difficult to say, but evidently there is considerable difference."

That the same child varies in talkativeness from situation to situation was demonstrated by Nice (67) who recorded for her daughter all that she said for 13 hours. The mean vocal output was 13.4 words per minute and the range was 7.4 to 22.2 words. Concerning her record Nice stated: "The hour she talked most was from ten to eleven when she used 1,333 words when playing with a three-year-old boy." Variation with setting is shown by the fact that she spoke 1,061 words when outdoors playing with children, 925 words at dinner, 902 words upon first getting up in the morning, and 444 words when on a walk.

Variation with change in setting in the nursery school was noted by Goodenough (32) who obtained a correlation of .42 (or .59 according to the Spearman-Brown formula) between the talkativeness scores at lunch and during the morning free-play hour.

8. *Differences in School Attendance*

It will be recalled that the subjects in the groups of the present study were selected without regard for the length of attendance at the nursery school. Inasmuch as this factor was not controlled in the present investigation it is of especial interest to know the findings of others relating to it.

The effect of school attendance was studied by Davis (16, p. 26) who found for the 248 kindergarten children who were five and one-half years of age that the coefficient of correlation (Pearson product-moment method) between talkativeness and length of school experience was .02. However, Goodenough (32) observed indications that "increased length of nursery school attendance makes for increased talkativeness in the nursery school situation during the first few months but thereafter little further change takes place."

Observation has indicated that the social adjustment of the child tends to affect his verbal output, therefore two studies relating to the influence of the nursery school upon adjustment are relevant.

Malley (61) found that 21 children observed for five-minute periods during the first three weeks at nursery school, and then six months later, showed: (*a*) more time in social and group contacts; (*b*) increase in use of techniques eliciting successful contacts; and (*c*) increase in number of children contacted. The findings of Hattwick (37) with 106 preschool children were similar. These subjects were matched for age, sex, nationality, race, and economic level. The group which had been in nursery school for nine months surpassed the six weeks' group in social adjustments and in behavior indicative of fewer inhibitions.

9. *Differences in Number of Siblings*

In a study of the only child Davis (18) observed that 15.4 per cent more of the only girls were rated as very talkative as compared with non-only girls. Greater talkativeness was present in 7.1 per cent more of the only boys than non-only boys. However, Goodenough (32) found in her investigation that "talkativeness appears to bear no significant relationship to size of family or position of child among siblings."

B. ANALYSIS OF DATA

The data in this chapter are based upon analysis of the uniform records (Table 4). As previously described, for each of the subjects four records from each setting were utilized. This was made possible by obtaining the means for four Outdoor periods, for instance, by averaging Periods 1, 5, and 9; Periods 3, 7, and 11; Periods 2, 6, and 10; and Periods 4, 8, and 12.

Only comprehensible verbal behavior is treated in the present section, since it seems inadvisable to put meaningless syllables on a par with words. The incidence of non-comprehensible vocalization will be given in a later paper.

1. *Reliability of Language Observations*

Smith (78, p. 17) stated that the variability in the amount of language used by the same child was too great for the number of words spoken to be "an adequate criterion." In this connection it should be noted that she observed most of the 88 cases which were used for that portion of the investigation for a one-hour period and only 25 of the subjects were seen more than once.

In the present study each child was observed for six hours, including miscellaneous observations, in periods of 10 minutes each. The periods extended over approximately 15 days, the mean time from obtaining the first record in the case of the uniform series, or 21 days where miscellaneous records were involved. When many language records are obtained, as in the present investigation, with a total of 36 observations, or 28 uniform observations, per child, the amount of comprehensible vocalization should become a fairly reliable criterion of language development. Robinson and Conrad (75, p. 164) estimated that "to secure a satisfactory, stable average of talkativeness of each child, with age held constant, would require about five hours of observation per child, distributed over many days."

In securing the reliability coefficients for the age groups as shown in Table 13 only the Outdoor and Indoor periods have been used since these seem most clearly to resemble those environments employed in Smith's study. The Pearson product-moment correlation technique was used to obtain correlations between the odd-even periods for the Outdoor and Indoor settings. These two situations constitute twelve and eight, respectively, of the 28 uniform periods, as shown in Table 4. The data referred to in Table 13 were divided by the odd-even method, and since this method takes account of only one-half the actual data, the coefficients were corrected by using the Spearman-Brown prophecy formula. Although it has been asserted by some writers that the Spearman-Brown formula over-corrects, it is used here and in all other instances involving reliability coefficients since it seems to be the best approach to the problem. The findings are shown in Table 13. The results indicate satisfactory reliability, with the correlations ranging from .84±.019 at the age of 30 months to .90±.013 at 48 months. The amount of comprehensible verbal behavior recorded per 10-minute period appears to be a reasonably reliable measure of language development in the present study.

TABLE 13

RELIABILITY COEFFICIENTS OF NUMBER OF COMPREHENSIBLE WORDS SPOKEN PER 10-MINUTE PERIOD FOR AGE GROUPS (BASED ON OUTDOOR-INDOOR UNIFORM PERIODS)

CA	Pearson *r*	*PE*
30	.84	±.019
36	.87	.018
42	.89	.014
48	.90	.013
54	.88	.013
60	.85	.023

2. *Differences Accompanying Changes in Setting*

What differences in vocal output, stated in terms of the number of comprehensible words per 10-minute period, are concomitant with changes in setting? It has been mentioned in a previous chapter that probably certain factors were operative in at least two of the settings to reduce language. The Indoor setting usually came immediately after the subject's arrival at school, when the influence of parting from his parents, need for time to adjust for the day and possibly, sleepiness, were factors of importance. In addition, favorite playmates sometimes had not yet arrived or were playing in another room, thereby depriving the child of an effective stimulus for vocalization.

As was stated previously, the Dinner setting, with the arbitrary selection of the persons who should be at the table with the subject, and his preoccupation with eating, seemingly would tend to inhibit speech. The Outdoor setting had a minimum of restrictions. In the Picture setting attendants actually elicited responses by asking questions if the subject did not talk spontaneously.

In Table 14 may be found the "Central Tendencies and Deviations for Number of Comprehensible Words per Subject per 10-minute Period, for Regular Boys and Relief Boys," according to setting. The means for Regular and Relief boys in all settings are 46.91 and 40.02 words, respectively. The results differ from setting to setting. The comprehensible vocal output is largest for both groups in the Picture setting and least in the Dinner situation. The respective means for Regular and Relief boys in the Picture setting are 56.35 and 42.80 words, and for the Dinner situation, 39.20 and 36.61 words.

TABLE 14

CENTRAL TENDENCIES AND DEVIATIONS FOR NUMBER OF COMPREHENSIBLE WORDS PER SUBJECT PER 10-MINUTE PERIOD, FOR REGULAR BOYS AND RELIEF BOYS

Setting	*Regular boys* Mean	*SD*	SD_m	*Relief boys* Mean	*SD*	SD_m
Out.	49.23	19.47	2.18	41.38	16.20	1.81
In.	42.85	17.10	1.91	39.29	14.24	1.59
Din.	39.20	14.59	1.63	36.61	14.69	1.64
Pict.	56.35	21.75	2.43	42.80	14.50	1.62

In Appendix *B* may be found, for Regular and Relief boys, correlations between the several combinations of settings, as, between the Outdoor-Indoor situations. It will be observed that the 12 correlations are .90 or higher, with probable errors ranging from .005 to .013. The subjects who talk a great deal in one setting also have records of volubility in the other settings, while those who are taciturn in one situation tend to maintain that characteristic in the other situations. The correlations indicate marked consistency in the amounts of verbal behavior recorded for the subjects in the several settings.

The mean number of comprehensible words per 10-minute period in the four different settings for Regular and Relief girls may be seen in Table 15. The means for the combined settings are 60.77 and 47.27 words, respectively. The Outdoor situation is the scene of the greatest amount of speech on the part of Regular girls, the mean being 75.13 words. Relief girls resemble boys in having the largest output, 52.66 words, in the Picture setting. For both groups of girls, as with boys, the smallest amount of speech occurs in the Dinner setting. Correlations between the several combinations of settings, as given in Appendix *B*, do not fall below .90, with probable errors ranging from .007 to .016. As was the case for boys, the girls who are loquacious in one setting tend to show that trait in the other situations, whereas the girls who say little in one setting tend to demonstrate the same characteristic in the other environmental situations. The correlations indicate marked consistency in the records obtained in the several settings.

TABLE 15

CENTRAL TENDENCIES AND DEVIATIONS FOR NUMBER OF COMPREHENSIBLE WORDS PER SUBJECT PER 10-MINUTE PERIOD, FOR REGULAR GIRLS AND RELIEF GIRLS

	Regular girls			Relief girls		
Setting	Mean	*SD*	SD_m	Mean	*SD*	SD_m
Out.	75.13	24.91	3.02	48.53	17.06	2.07
In.	59.68	20.15	2.44	45.52	15.46	1.87
Din.	47.22	15.23	1.85	42.37	16.25	1.97
Pict.	61.06	21.30	2.58	52.66	17.76	2.15

Regular subjects, as shown in Tables 14 and 15 show greater variability than do Relief cases, except in the Dinner setting. Regular girls are the most variable and Relief boys the least variable of the four groups. Regular girls and Relief girls are more variable than the corresponding groups of boys of approximately the same socio-economic levels.

In Table 16 are given "Critical Ratios of the Differences between Specified Settings in the Number of Comprehensible Words per 10-Minute Period for the Groups Indicated." The correlations involved in obtaining the critical ratios may be found in Appendix *B*, and the means, standard deviations, and standard errors of the means are given in Tables 14 and 15. The standard error of the mean was used as the measure of reliability and a critical ratio of 3.00 may be taken as indicative of a truly significant difference, although one slightly less than this does not necessarily indicate non-significance. In giving comparisons of settings, the situation which has the larger mean will be mentioned first.

In the self-comparisons of Regular boys in the several settings, all differences satisfy the criterion of statistical reliability and the smallest critical ratio is 6.40. The environments which were the scenes of the greatest amounts of speech are the first-named of each comparison: Picture-Indoor; Picture-Dinner; Picture-Outdoor; Indoor-Dinner; Outdoor-Indoor; and Outdoor-Dinner.

For Relief boys (Table 16), five completely reliable differences are the result of the self-comparisons of records in the four settings, as follows: Picture-Indoor; Picture-Dinner; Outdoor-Indoor; Outdoor-Dinner; and Indoor-Dinner. Although it does not indicate complete statistical reliability the critical ratio of 2.73 for the Picture-Outdoor settings indicates approximately 99.70 chances in 100 that the true difference is greater than zero.

TABLE 16

CRITICAL RATIOS OF THE DIFFERENCES BETWEEN SPECIFIED SETTINGS IN THE NUMBER OF COMPREHENSIBLE WORDS PER 10-MINUTE PERIOD FOR THE GROUPS INDICATED (CORRELATED OBSERVATIONS)

Setting	In.	Pict.	Din.	Setting	In.	Pict.	Din.
	Regular boys				*Relief boys*		
Pict.	12.27			Pict.	5.16		
Din.	6.40	15.18		Din.	4.79	9.52	
Out.	11.19	7.92	12.38	Out.	3.12	2.73	6.72
	Regular girls				Relief girls		
Pict.	1.92			Pict.	7.93		
Din.	11.12	12.14		Din.	4.32	12.25	
Out.	13.92	11.73	18.99	Out.	4.63	4.59	8.11

In the self-comparisons of Regular girls (Table 16) in the several settings only one difference, that for the Picture-Indoor settings, is unreliable. Five of the six critical ratios are indicative of highly significant differences between the settings which have been compared, as follows: Outdoor-Dinner; Outdoor-Indoor; Outdoor-Picture; Indoor-Dinner; and Picture-Dinner.

The amount of language used by Relief girls likewise varies from setting to setting. All differences are thoroughly reliable statistically, and the larger means occur in the first-named settings in the following comparisons: Picture-Indoor; Picture-Dinner; Picture-Outdoor; Outdoor-Indoor; Outdoor-Dinner; and Indoor-Dinner.

If the observations are treated as if they are uncorrelated, the data in Table 16 are changed markedly. For instance, for Regular boys only the following three critical ratios satisfy the criterion for statistical significance: Outdoor-Dinner; Picture-Indoor; and Picture-Dinner. For the Outdoor-Indoor and Picture-Outdoor comparisons the respective critical ratios of 2.20 and 2.18 indicate approximately 98.6 chances in 100 of a true difference. For Relief boys, the treatment of the data as if uncorrelated produces no significant differences between settings, whereas five resulted from the formula involving the correlations. The largest critical ratio, 2.69, occurs for the Picture-Dinner settings, and indicates about 99.7 chances in 100 that the difference is a true one. Regular girls with either procedure have five out of six critical ratios which indicate statistical reliability. When the correlated observations are used, the critical ratios are much larger. Treatment of the observations of the records of Relief girls as if they are uncorrelated reduces the number of significant differences from six to one, the Picture-Dinner comparison. However, the Outdoor-Dinner, and Picture-Indoor settings have respective critical ratios of 2.15 and 2.51, indicating approximately 98.3 and 99.4 chances in 100 that the true differences are greater than zero.

In Table 17 may be found the critical ratios of the differences between settings for "All Subjects." In this case the method of computing the critical ratios was not the same as that employed in connection with the ratios given in Table 16, inasmuch as the ratios for "All Subjects" (Table 17) are not based upon a formula which involves the correlations between settings. Undoubtedly, the critical ratios for "All Subjects" would have been larger if the same procedure as was used for the results in Table 16 had been followed. Even with the method which was employed, the critical ratios are indicative of significant differences except in the case of the Picture-Outdoor setting, and in that instance the result is zero. For three of the four individual groups the differences between these two settings (Picture-Outdoor) are statistically reliable, but only in two instances were the records of the former significantly superior to those in the Outdoor setting. For the combined groups of "All

Subjects" the larger means may be noted in the several comparisons, since the situation in which the greater amount of language occurred will be given first: Outdoor-Indoor; Outdoor-Dinner; Indoor-Dinner; Picture-Indoor; and Picture-Dinner.

From the data just described, it is clear that the four groups apparently are affected by differences in the immediate setting, as that term is described in the present investigation, or that changes in the amount of verbal behavior are concomitant with changes in the setting. Inasmuch as 22 of the 24 self-comparisons of specific groups, and five of the six comparisons for "All Subjects" in the several settings result in highly significant differences, it seems reasonable to conclude that the factor of setting should be carefully controlled when securing samples for a study of the amount of comprehensible speech. Since reliable differences between settings occur, apparently it is advisable, if representative samples are desired, that they be secured in several different types of situation. If this is not possible the setting which is used should be described in detail and the limitations recognized.

That the amount of comprehensible speech used in a given unit of time may be a function of the setting is an hypothesis for which ample evidence has been secured. The complex of factors composing the settings may function in a unified fashion, or it is possible that single variables in the situations are responsible for the differences that obtain. For the groups involved, under the conditions described, it appears that certain differences between settings may account for statistically significant differences in the output of comprehensible speech, or that changes in the amount of vocalization are concomitant with changes in setting.

3. *Variations Accompanying Differences in Socio-Economic Status or Sex*

Table 17 shows the "Critical Ratios of the Differences between Compared Groups in Mean Number of Comprehensible Words per 10-Minute Period, in the Settings Specified." Critical ratios for group comparisons in Table 17 differ from those in Table 16 in that the former are based upon uncorrelated observations, inasmuch as different groups are being compared. Comparisons of the amount of language used by Regular and Relief boys indicate statistically significant differences only for records made in the Picture setting,

TABLE 17

CRITICAL RATIOS OF THE DIFFERENCES BETWEEN COMPARED GROUPS IN MEAN NUMBER OF COMPREHENSIBLE WORDS PER 10-MINUTE PERIOD, IN THE SETTINGS SPECIFIED

Setting	Out.	In.	Din.	Pict.	All	Setting	Out.	In.	Din.	Pict.	All
Regular boys—Relief boys						Regular girls—Relief girls					
Out.	2.77					Out.	7.27				
In.		1.43				In.		4.61			
Din.			1.12			Din.			1.80		
Pict.				4.64		Pict.				2.50	
All					2.66	All					4.31
Regular boys—Regular girls						Relief boys—Relief girls					
Out.	6.94					Out.	2.60				
In.		5.43				In.		2.53			
Din.			3.25			Din.			2.25		
Pict.				1.33		Pict.				3.67	
All					4.40	All					2.83
Regular boys—Relief girls						Relief boys—Regular girls					
Out.	0.23					Out.	9.59				
In.		1.00				In.		7.01			
Din.			1.24			Din.			4.30		
Pict.				1.14		Pict.				5.99	
All					0.13	All					7.08
All subjects						Boys—Girls					
Out.	——	3.82	7.30	0.00	——	All					3.68
In.			3.73	4.15	——	Regular—Relief					
Din.				7.99	——	All					3.58

but the critical ratio of 2.77 for Outdoor records gives approximately 99.74 chances in 100 of a true difference. There is not a significant difference between the two groups when they are compared on the basis of the combined settings, although the critical ratio of 2.66 indicates about 99.7 chances in 100 that a true difference exists. In this group comparison the socio-economic status differs and the sex factor is constant.

When Regular girls and Regular boys are compared, the socio-economic status being relatively constant and sex differing, the Outdoor, Indoor, Dinner, and combined (All) records show significant differences. In the Picture setting where language was elicited, if necessary, by questions the critical ratio is 1.33, which means that there are approximately 90 chances in 100 of a true difference.

Regular girls used more words than did Regular boys in each of the four settings. The experimenter observed that boys seemed to engage in more strenuous activity than girls. This possibly inhibited the language output, especially in the Outdoor setting, and may have caused more errors in recording for boys. This statement is based, however, only upon casual observation and not upon careful activity records.

The group comparisons show Relief girls and Regular boys to be more nearly alike than any groups in the number of words spoken per 10-minute period. The largest critical ratio, 1.24, is for the Dinner setting, and the critical ratio for the combined settings is only 0.13. The similarity between the two groups is particularly noteworthy since they differ both in sex and socio-economic status.

The records of Regular and Relief girls, with a difference in socio-economic levels to be considered, show statistically significant differences for Outdoor and Indoor settings. In these situations the critical ratios are 7.27 and 4.61, respectively. A critical ratio of 2.50 in the Picture situation indicates approximately 99.4 chances in 100 that the true difference is greater than zero. Regular girls have a larger amount of verbal behavior in all situations, and the critical ratio of 4.31 for the combined settings indicates the existence of a reliable difference between the two groups. The experimenter did not notice any difference in the amount of physical activity engaged in by these two groups, as was the case for boys, although some may have existed as a possible explanatory factor.

It is of interest to observe the difference between Relief girls and Relief boys, groups of relatively the same socio-economic status, but differing in sex. Relief girls in each case used more language, but all critical ratios indicate unreliable differences except for the Picture setting. However, the chances are 98.8 in 100, or greater, that the true differences are more than zero for comparisons made on the basis of the Outdoor, Indoor, and Dinner settings. For the combined settings the critical ratio is 2.83, which gives approximately 99.8 chances in 100 of the existence of a true difference in favor of Relief girls when compared with Relief boys.

The largest differences between any two groups are those resulting from comparisons of Regular girls and Relief boys. Regular girls are decidedly more talkative and all critical ratios indicate highly significant differences. For the combined settings the critical ratio is 7.08. The greatest difference is for the Outdoor setting, in which Regular girls have an average of 75.13 words per 10-minute period as compared with 41.38 words for Relief boys. In this comparison of groups both sex and socio-economic differences are present. As previously stated in connection with another comparison, it appeared to the experimenter that boys tended to engage in more vigorous activity in the Outdoor setting and this may have reduced the language output and caused more errors in recording the speech of Relief boys.

Are the combined groups of girls more talkative than the two groups of boys? The critical ratio is 3.68 (Table 17), a figure which indicates a difference which is statistically reliable, in favor of girls. This fact is worthy of note because of the conflicting findings on sex differences obtained by other investigators, as mentioned previously. The differences between the two groups of girls and Relief boys are marked, there being five significant differences in the amount of language used in specific settings as compared with three such differences between settings from the comparison of the two groups of girls and Regular boys. On the basis of the combined settings (All), critical ratios indicate significant differences when Regular girls and Regular boys are compared, but show a surprising similarity, as revealed by a critical ratio of 0.13, between the records of Relief girls and Regular boys.

Are the combined groups of Regular subjects more talkative than the two groups of Relief cases? The critical ratio of 3.58 (Table 17) is indicative of a difference which is statistically reliable, in favor of Regular subjects. As previously indicated in connection with sex differences, the superiority of Regular subjects is much clearer in comparison with Relief boys than Relief girls. There are five significant differences between Regular subjects and Relief boys in the amount of language used in the specific settings, and only two reliable differences between Regular subjects and Relief girls.

It is of interest to observe the instances where sex differs and the socio-economic factor is relatively constant. This is true of Regular girls-Regular boys and the critical ratio is indicative of a truly significant difference on the basis of the combined settings. A similar type of grouping is found for Relief girls-Relief boys, and the results indicate approximately 99.8 chances in 100 that a true difference exists.

It is likewise important to note the instances in which the socio-economic factor varies and the sex factor is constant, as for Regular boys-Relief boys. Comparisons made on the basis of the combined settings indicate about 99.7 chances in 100 that there is a difference greater than zero which is not due to chance. A similar type of grouping is found for Regular girls-Relief girls, and the differences between two of the four settings are statistically reliable, as is also the critical ratio for the combined settings.

Although both factors appear to be influential in specific comparisons, the effects of sex and socio-economic status do not entirely explain the differences existent in the amount of verbal behavior. The findings are not equally indicative of positive relationships. In this connection, the fact should be restated that the critical ratios for the comparisons in the combined settings of Regular girls-Relief boys, and Relief girls-Regular boys, with sex and socio-economic differences existent in both pairs, are 7.08 and 0.13, respectively. Neither sex nor socio-economic status can be regarded as the sole explanation of differences in vocal output when the comparisons of similar types of groupings result in such contradictory findings.

The experimenter found some evidence in the present study of the difference between socio-economic groups described by Gesell and Lord (30). In the present investigation the reduced vocal output was most characteristic of Relief boys. It will be recalled, as stated in a previous chapter, that the Relief cases came from homes of relatively low socio-economic status and, in addition, from homes which were especially affected by the depression. Several observers who saw the Nursery School with Relief and Regular subjects commented upon the difference in what Gesell and Lord have termed "spontaneity." Although not all Relief subjects showed a lack of that attribute, the differences between specific cases of the two socio-economic groups were noticeable. Relief children were more serious or "prosaic," in accordance with the findings of Markey (62, p. 130) who described the imaginative play of children from the lower economic groups as compared with that of the more fortunate levels. The Relief children in the present study often played that they were ironing, sweeping, boiling clothes, drawing water from the well, cooking, and the like. Obviously, these activities were imitative of home activities. The Regular subjects tended to be less serious in dramatic play and there was not as much imitation of those adult activities which are classed as manual. Such characteristics of Regular and Relief subjects may account for some of the existing language differences between the two groups.

4. *Differences Accompanying Variations in Age*

Does the vocal output, stated in terms of the number of comprehensible words spoken per 10-minute period vary with age? In Table 18, which is based upon the Outdoor-Indoor uniform records, may be found the distribution by age, group, and sex. The data are confined to records of these settings, since they most nearly resemble those used by Smith (78). It will be observed that the means are not identical with those of Tables 14 and 15, inasmuch as the data in Table 18 are based on age groups, and Tables 14 and 15 disregard age groupings. In each of the four groups (Table 18), verbal behavior shows an increase with age. Regular girls and Relief girls of the 54 months' age level have the best records of the four different groups, with 84.9 and 75.2 words per 10-minute period, respectively. However, the record of Regular girls at 48 months is superior to that just mentioned for Relief girls at 54 months. At each age girls exceed boys in the amount of comprehensible verbal behavior. The poorer showing of the latter is due largely to the inferior records of Relief boys rather than to those of Regular boys. At all ages Regular subjects surpass Relief subjects.

TABLE 18

MEAN NUMBER OF COMPREHENSIBLE WORDS SPOKEN PER 10-MINUTE PERIOD BY AGE, GROUP, AND SEX, AND RESULTS OF SMITH'S STUDY (BASED ON OUTDOOR-INDOOR UNIFORM RECORDS)

Group	Boys			Girls				Smith's
(Mos.)	Reg.	Rel.	All	Reg.	Rel.	All	All	data
30	22.1	14.6	18.4	27.4	17.9	22.7	20.6	19.7
36	37.7	22.5	30.1	43.2	35.3	39.3	34.7	37.2
42	55.4	37.2	46.3	61.7	50.7	56.2	51.3	57.3
48	61.9	45.5	53.7	78.4	62.1	70.3	62.0	66.7
54	74.6	58.7	66.7	84.9	75.2	80.1	73.4	69.2
60	50.4	43.9	47.2	72.1	64.6	68.4	57.8	66.7

The range of the means for Regular boys is from 22.1 words per 10-minute period at the 30 months' level to 74.6 words in that period of time for the 54 months' age group. The range, 14.6 to 58.7, is smaller for Relief boys at those age levels. The mean for

Regular girls is 27.4 words at 30 months and 84.9 words at 54 months, while Relief girls average 17.9 and 75.2 words at those ages. The smaller records at 60 months may be due to an error in sampling, since there is a relatively small number of cases at that age level. In addition, it should be stated that usually not more than one or two children of that age group chanced to be in attendance at the Nursery School during a quarter. For a short time, attendance of children over 60 months of age was permitted but not encouraged, and the practice soon was entirely discontinued. It is possible that the lack of playmates of the same age, as well as other factors, may have influenced the reduction in verbal behavior for subjects of that age level.

The combined records of the four groups show a gradual increase in the amount of comprehensible vocalization per 10-minute period with increase in age, except at the 60 months' level. Smith (78, p. 17) used 124 records of 88 different children and observed each subject for an hour during a play period with other children. She (78, p. 27) concluded that "the average total number of words per hour increased with age." Smith's results, reduced to the number of words per 10-minute period, are given in Table 18, in connection with those of the present study. The subjects varied in number from 16 to 23 at the age levels used. It will be seen that the records of her cases somewhat excel those of the present investigation at four of the six age levels, and in two instances the means for the present study are larger. However, the differences between the means are relatively small at all ages, although the method of grouping the subjects for age levels was not identical in the two studies.

5. *Differences Accompanying Variations in Intelligence*

The present study indicates that significant differences exist between the amounts of comprehensible language of certain socioeconomic groups when group comparisons are made. These groups differ in intelligence as measured by the *Merrill-Palmer Scale of Mental Tests* (82). The comparisons, based on the critical ratios for the combined settings, will be found with the percentile ranks in Table 19. It should be noted that Regular subjects, with a percentile ranking of 70, surpass Relief subjects who have a mean percentile ranking of 46, and the difference is statistically reliable.

TABLE 19

Percentile Ranks Obtained on *The Merrill-Palmer Scale of Mental Tests*, Means for the Groups, and Critical Ratios of the Differences Between Groups for Number of Comprehensible Words Per 10-Minute Period*

Groups compared	Words per ten minutes		Percentile ranks		Critical ratios
Regular—Relief	53.84	43.65	70	46	3.58
Girls — Boys	54.02	43.47	58	58	3.68
Reg. B.—Rel. B.	46.91	40.02	69	46	2.66
Reg. G.—Reg. B.	60.77	46.91	71	69	4.40
Reg. G.—Rel. G.	60.77	47.27	71	45	4.31
Rel. G.—Rel. B.	47.27	40.02	45	46	2.83
Rel. G.—Reg. B.	47.27	46.91	45	69	0.13
Reg. G.—Rel. B.	60.77	40.02	71	46	7.08

*In naming the groups and in stating means and percentile ranks, the results of the superior group are given first.

Likewise, significant or nearly significant differences occur when Regular boys-Relief boys, and Regular girls-Relief boys are compared, the percentile ranks being 69-46, and 71-46, respectively. However, a statistically insignificant difference results when Relief girls and Regular boys, with percentile ranks of 45 and 69, are compared. Actually, Relief girls slightly surpass Regular boys in the amount of comprehensible verbal behavior. In addition, it should be noted that reliable differences occur for groups having the same, or practically the same, percentile ranks, as in the comparison of Girls-Boys, and Regular girls-Regular boys.

6. *Differences Accompanying Variations in Number of Siblings*

The only child has been thought by some authorities to occupy a fortunate position, with greater opportunity for adult contacts and therefore more chance to learn language. Other writers consider the only child to be handicapped by the lack of siblings. The findings of Goodenough (32) and Davis (18) on this subject have been cited earlier. Goodenough and Leahy (35) found that "only children in the kindergarten groups are rated as more aggressive and more self-confident" than any of the other subjects. They are also "highly gregarious in their social interests." Since one would expect these traits to correlate positively with the amount of language used, it is of interest to observe the percentage of only children in the several groups. The criterion of an only child in the present study is a record of no siblings less than 15 years of age at the time of the language recording. The incidence of only children in the four groups is as follows: Regular boys, 55 per cent; Regular girls, 53 per cent; Relief boys, 10 per cent; and Relief girls, six per cent. Regular subjects, with far greater numbers of only children in their groups, are significantly superior to Relief subjects in number of words spoken. It cannot be said, however, that the fact of being an only child is in itself an important cause of language superiority. Again, the close resemblance between the records of Regular boys and Relief girls, two groups whose records of "onliness" were 55 per cent and six per cent, respectively, must be considered. Regular girls are significantly superior to Regular boys in amount of language spoken, yet they have almost the same records for "onliness." Relief girls markedly surpass Relief boys in vocal output, the critical ratio of 2.83 indicating approximately 99.8 chances in 100 of a true difference, yet the incidence of only children is very similar in the two groups.

The number of siblings presents a similar problem worthy of mention when studying verbal behavior. The four groups had the following records of siblings who were less than 15 years of age at the time of the language recording for a specific subject: Regular boys, 12; Regular girls, 9; Relief boys, 31; and Relief girls, 23. Relief subjects, with many more siblings, are significantly inferior in comprehensible verbal output. Again, however, it should be noted that the speech records of Relief girls-Regular boys are very similar although they have totals of 23 and 12 siblings, respectively. Also should be mentioned the statistically reliable difference between Regular girls-Regular boys, groups having a similar number of siblings per child. This factor may affect the amount of language to some extent, but in the present investigation its possible influence is not consistent.

7. *Differences Accompanying Variation in Other Factors*

The most unusual problem at any age level was that of a Relief boy in the 42 months age group. The case is of interest inasmuch as his was the most serious instance of maladjustment encountered over a five-year period in the Nursery School. Six months prior to the making of his language records, observers who watched him closely reported that during two hours of daily observation for a week he had not spoken a word or used any vocal expressions. Occasional shaking of the head was noted, but no other gestures. Less careful observation at other times also indicated refusal or inability to use language. Every effort was made to help the child to adjust, and the members of the staff and the student observers coöperated to develop his language. His improvement in willingness and ability to speak was slow, and was associated with increased ability to play with other children. An achievement of eight words per 10-minute period after approximately six months was particularly commendable in his case, although still far below the average, which was 37.2 words (Table 18), for his age and socio-economic group.

There are marked individual differences in the amount of language used. Certain variations occur which apparently are influenced by, but may merely be the concomitants of, changes in setting, age, sex, socio-economic status, and number of siblings. Differences due to emotional factors also appear to be important. Unless evidence which is not now available is presented, it may be assumed that individuals are distributed somewhat according to the normal probability curve in the amount of vocalization. The implications of such an assumption have been clearly stated by Edwards (24). Judgments based on extreme cases in which unusual records of verbal behavior are regarded as typical are to be avoided.

C. SUMMARY

1. The reliability coefficients, obtained by correlating the odd-even-numbered records of the amount of comprehensible language used in the uniform Outdoor and Indoor settings, are .84±.019, or higher, for the several age groups.

2. Correlations between any two of the settings are .90 or more, indicating a high degree of consistency in the amount of verbal behavior recorded for the subjects in the four settings.

3. The comprehensible words spoken per 10-minute period in the combined settings are as follows for Regular boys, Relief boys, Regular girls, and Relief girls, respectively: 46.91, 40.02, 60.77, and 47.27.

4. The amount of comprehensible speech shows variations which apparently are concomitant with changes in setting, as follows:

(*a*). Of the 24 self-comparisons of specific groups in the several settings 22 result is highly significant differences.

(*b*). Of the six self-comparisons in the several settings for

"All Subjects" as a combined group, five result in completely reliable differences. These facts indicate the need for controlling the factor of setting when securing language samples.

5. The amount of comprehensible speech shows variations which apparently are concomitant with changes in socio-economic status or difference in sex, as follows:

(*a*). Girls are superior to boys, and Regular subjects to Relief subjects, the differences being statistically reliable.

(*b*). Where sex differs and the socio-economic factor is relatively constant, or where the socio-economic status differs and the sex factor is constant, the differences between groups are completely reliable, or closely approach statistical significance. However, the results are not consistent when both sex and socio-economic differences are present in the comparisons.

6. The amount of comprehensible speech shows variations which apparently are concomitant with changes in age, as follows:

(*a*). With increase in age the amount of speech increases, up to the age level of 60 months. At all ages girls excel boys, and Regular subjects excel Relief subjects. Relief boys have the smallest records at all age levels.

7. The amount of comprehensible speech shows variations which apparently are concomitant with differences in intelligence or variations in the number of siblings, as follows:

(*a*). *Regular subjects, with a percentile rank of 70, are significantly* superior to Relief subjects whose percentile rank is 46. However, there are several inconsistencies from other group comparisons which indicate that intelligence as a possible causal factor does not exert a uniform influence in all instances.

(*b*). Regular groups, composed of a much larger group of only children and having few siblings as compared with Relief subjects, significantly surpass the latter in verbal behavior. Nevertheless, the inconsistencies which may be noted from a study of other group comparisons indicate that sibling incidence as a possible causal factor does not exert a uniform influence in all instances.

8. Individual records reveal that certain variations in the amount of language may be attributed to, or be concomitant with, serious maladjustments.

V. PARTS OF SPEECH

A. RELATED STUDIES

1. *Difficulties Involved*

Jespersen (45, p. 66) stated that in working with linguistic subjects "it is necessary to have names for the various classes into which words fall naturally, and which are generally, but not very felicitously, called 'parts of speech.'" This expression is used, therefore, in accordance with his terminology.

The classification of children's language into the conventional forms of grammar has been attempted by various investigators although difficulties and differences of opinion are numerous. Representative of one viewpoint is the statement by Markey (63, p. 84) that "any such classification of a child's speech into the conventional forms of grammar is bound to be forced." Dewey (21) believed that the "psychological classification is to class the word according to what it means to a child, not to the adult with his grammatical forms all differentiated." McCarthy (57, p. 126) analyzed the general problem into specific problems, as follows: "The question of the part of speech under which the single-word sentence is to be classified, the problem of the inclusion of variants of words, the use of the total number of words or simply of the number of different words used—all make material of this sort elusive."

Further disagreements are reflected when the actual classifications of the parts of speech are examined. As stated in a previous chapter, many of the divergencies in results are due to the different definitions for the parts of speech held by authorities on grammar.

Inasmuch as the reader who has not made a study of the field may not be cognizant of the differences involved, attention is called to the fact that counts based upon the number of different words used and counts based upon the total number of words used may give quite different results. McCarthy (57, pp. 114 and 118) and Zyve (92) classified data according to the two methods. Numerous vocabulary studies are available, but they will merely be mentioned here since the present investigation treats only the total number of words used by the subjects. Bateman (4) gave the vocabulary count for his two children. Mateer (64) summarized several earlier studies. Waddle (86, p. 167) tabulated the findings from 46 investigations. Drever (22) presented the data for three subjects. Tracy and Stimpfl (84) gave tables of the vocabularies of 12 children. Grant (36) summarized numerous vocabulary studies. Descoeudres (20) gave separate norms for children of the upper and lower occupational groups in connection with her vocabulary study. Horn (39) made a thorough investigation of the most common words in the spoken vocabulary of preschool children, including subjects in the sixth year. Williams and McFarland (91) discussed various aspects of vocabulary development in connection with *A Revision of the Smith Vocabulary Test for Preschool Children.* Many of the references indicated for the study of the total number of words also give data for the number of different words.

2. *Sex Differences*

In studying the total number of words used, McCarthy (57, p. 114) found that the greatest differences between the sexes occurred at the earliest age levels, 18 and 24 months. Differences thereafter were usually very small. At 18 months, girls excelled boys in nouns by approximately eight per cent. McCarthy stated that at this age "this relationship probably indicates a more advanced stage for the girls than for the boys." Also, girls used more adjectives and adverbs than did boys at the two early levels. Girls began the use of conjunctions before boys, but the latter surpassed in the use of interjections, especially in the first age group.

Stern (81, p. 154) believed that "the quick reaction to what has been heard with corresponding utterances is, in proportion, most frequently found in girls, whilst the longer storing up of impressions and only indirect imitation is more common in boys." He (81, p. 161) also noted a tendency for the language of boys to show more spontaneity, remarking that "effervescing self creations" or "clumsy originality" scarcely ever seemed to occur in girls.

Davis (16, p. 117), using subjects of five and one-half, six and one-half, and nine and one-half years of age, found that the use of personal pronouns was greater for girls than for boys and the use of conjunctions was slightly greater for girls.

3. *Socio-Economic Differences*

A restricted environment, which may be assumed to be more characteristic of the child from the lower occupational groups than of the child from the more favored classes, was considered by Drever

(22, p. 102) to cause a decrease in the incidence of nouns. McCarthy (57, p. 123) tested the hypothesis by investigating the parts of speech used by the upper and lower occupational groups. Her data indicated a larger proportion of nouns and adjectives occurring in the language of the latter subjects. The upper group surpassed the lower, particularly in verbs and pronouns, and with smaller differences, in adverbs and conjunctions.

Van Alstyne (85, p. 57) suggested the possibility that vocabulary was slightly more related to environmental factors than were intelligence test scores. She found the correlation between vocabulary and environmental factors to be .70±.04, and the correlation between intelligence test scores and environmental factors to be .61±.05.

4. *Differences in Setting*

Boyd (8) indicated that the "child uses different words and frames different kinds of sentences with every change in environment." Davis (16, p. 108) found a similarity of responses in a setting that was carefully standardized. She reported that the "first two records yielded 204 different words, and throughout the analysis it was found that these recurred with astonishing frequency." Nice (70) noted the effect of setting and observed that the situation of preschool and day-nursery children evidently was a very different one from that of a child with her parents, judging from her results and from those of other investigators.

In the section on *Number of Words Spoken* several studies indicating differences in the vocal output in different settings were cited. Particularly, the findings of Nice (67), Ellesor (25), and Mattson, as reported by Markey (63), should be recalled. These writers noted the varying amounts of verbal behavior in different situations, but they did not investigate the possible effects upon the parts of speech.

McCarthy (56) obtained indirect evidence of the change in the type of words in a comparison of the responses of 31 children on the playground and in the experimental room with an adult. The emotionally-toned responses represented approximately nine per cent of all the responses in the latter setting, but 33 per cent of the remarks on the playground. This type of response included wishes, requests, threats, and commands, and would involve an increase in the incidence of verbs.

Smith (79, p. 188) called attention to the high incidence of questions and answers in McCarthy's situation in which the child was with adults. She also noted in her own study the higher frequency of imperatives on the playground, but more evidence of connected discourse and narration when the child was with grown-ups. The effect of the two types of situation upon the parts of speech was not specifically investigated, however, by McCarthy or by Smith.

5. *Age Differences*

It is an accepted fact that the use of appropriate pictures may stimulate responses which show age differences. Binet used this type of situation as one of his tests of intelligence, and Terman (83) utilized it in the Stanford Revision. Binet considered the child of three or four years as being in the "identification" stage when presented with such stimuli. Terman (83, p. 146) stated that "description is rarely encountered before five years, and interpretation rarely before nine or ten." The same pictures when used at years seven and twelve are expected to elicit description and interpretation, respectively. Stern (81, p. 206) regarded the third year as the "substance stage," and the seventh year as the "action stage." Apparently, the different types of responses at the several age levels would involve differences in the parts of speech, the "identification" stage being associated especially with a high incidence of nouns.

Nice (69) emphasized changes in language with age, basing her observations mainly upon a thousand or more words collected for her daughters at 30 months and at the ages of three, four, five, eight, and ten years. She stated that "most children at 30 months are in the short-sentence stage." At that time there is a "preponderance of nouns, a lack of articles, auxiliary and copulative verbs, prepositions, and conjunctions; the same is true of the short sentence to a lesser degree." The established sentence she found to be characterized by "a greater definiteness and complexity as is shown by the increasing use of these relational words that were largely lacking before." She estimated that "at three years perhaps half the children have passed into the established sentence stage, and only those retarded in speech have not attained it by four."

6. *Summary of Findings from Six Related Studies*

In Table 20 is given a summary of the findings of six investigators who reported the per cent of each part of speech used by preschool subjects. Boyd (8) obtained 1,250 unselected sentences near the birthdays of his daughter for ages two through eight. He also gathered 100 conversational sentences from each of ten men and eight women novelists. Day (19) studied 80 pairs of twins and followed McCarthy's technique of showing toys and pictures to the child. Using this procedure, she obtained 50 responses for each subject. As shown in Table 20, there were 40 subjects at each age. Inasmuch as a table with exact percentages was not given in Day's report of her findings, the data in Table 20 represent estimates based on Figure 7 of Day's investigation. McCarthy (57, p. 33) used 140 subjects and obtained a record of 50 remarks while the child was being shown toys or pictures. In this study she used subjects so selected that the age groups were representative of the percentages of the various occupations in the locality. Kirkpatrick (52, p. 52) gave the percentages for one child as secured "during an hour or more taken at regular intervals during the ages of two and four." Nice (67) recorded an hour's conversation for her daughters at ages three, four, and five. She had found that "with the total words, one hour gives a good prophecy of the all day conversation." Smith (79) used 220 subjects for whom 305 records were obtained. The records of cases more than 65 months of age are omitted from Table 20. From most of the subjects one hour of spontaneous conversation was secured, the records being made by the experimenter without questioning the subjects. In Smith's investigation the age classification of 18 months, for instance, includes subjects of 18-29 months, inclusive. It should be noted that the age groupings are not the same in the several studies, and not in all cases were the records obtained near the birthdays of the subjects. In making comparisons, the observer should recognize this difference as well as the variations in method of securing the data.

Since eminent grammarians do not agree on the problem of classifying parts of speech, and psychologists probably are influenced by authorities who hold opinions which are at variance, it is to be expected that studies of language development will show different results due to this factor. The fact that different grammars may have been used as the guides to classification should be considered in comparing such data as are found in Table 20. Nice (69, p. 243) recognized the variations in results which apparently could be attributed to this cause. She noticed that Boyd (7), who enumerates

TABLE 20

MEAN PER CENT OF EACH PART OF SPEECH REPORTED BY OTHER INVESTIGATORS (BASED ON TOTAL NUMBER OF WORDS USED)

Study by	Age	No. cases	Nouns	Pron.	Adj.*	Verbs	Adv.	Prep.	Conj.	Interj.	Misc.*
Boyd	24	1	36.8	6.2	13.7	28.2	13.2	1.4	0.2	0.3	
	36	1	16.4	15.8	17.4	29.6	9.4	7.9	3.4	0.2	
	48	1	14.8	18.1	14.6	30.0	11.6	7.6	2.9	0.3	
	60	1	14.7	18.7	15.5	29.6	10.1	6.9	4.0	0.4	
	72	1	15.3	17.6	15.8	28.6	10.1	7.7	4.4	0.4	
	84	1	14.3	18.5	15.4	28.2	10.8	7.3	5.0	0.5	
	96	1	14.6	18.4	15.2	28.3	10.2	7.7	5.2	0.3	
	Women	8	15.1	17.9	16.1	24.5	9.6	9.0	6.8	1.1	
	Men	10	16.5	17.2	16.7	24.2	9.2	9.2	6.6	0.6	
*Day***	24	40	35.0	8.0	2.0	14.5	9.5	2.0	0.0	16.5	12.5
	36	40	23.0	18.0	5.0	23.0	11.8	5.2	0.9	3.5	10.0
	48	40	22.0	18.3	6.0	23.5	10.0	6.2	2.0	3.0	8.5
	60	40	22.5	17.5	8.0	23.5	9.5	5.2	3.3	2.0	9.0
Kirkpatrick	28	1	38.7	5.1	19.5	22.5	8.5	3.8	0.0	1.2	
	34	1	18.0	27.8	5.9	33.6	7.8	5.6	0.2	0.0	
	46	1	15.4	26.5	7.0	31.0	13.4	3.8	2.7	0.7	
McCarthy	18	20	50.0	10.3	9.6	13.9	7.9	0.0	0.5	7.6	0.3
	24	20	38.6	14.6	10.3	21.0	7.1	3.6	0.5	2.4	1.8
	30	20	25.8	19.0	14.3	23.4	6.7	4.6	1.7	2.8	1.8
	36	20	23.4	19.2	16.1	23.0	7.0	6.9	2.4	1.5	0.5
	42	20	18.5	20.3	15.7	26.0	7.8	6.3	2.3	2.0	0.8
	48	20	20.1	21.6	14.6	26.0	5.9	6.7	3.6	0.8	0.8
	54	20	19.3	20.5	15.2	25.1	7.0	7.1	3.8	1.2	0.8
Nice	36	1	17.3	22.3	9.3	23.1	18.6	7.1	0.7	1.6	
	48	1	19.4	23.1	7.6	29.6	14.0	4.5	0.3	1.6	
	60	1	15.8	22.8	13.7	28.1	10.2	6.2	2.0	1.2	
Smith	18	65	28.0	13.0	5.0	23.0	18.0	1.8	0.3	10.2	
	30	75	18.0	23.0	9.0	26.0	16.0	4.4	1.6	5.5	
	42	75	16.0	24.0	11.0	26.0	12.0	6.5	2.2	2.1	
	54	75	15.0	24.0	13.0	27.0	12.0	6.4	2.5	1.8	

*Adjectives include articles. *Miscellaneous* for McCarthy's and Day's studies includes *yes, no, kinda,* etc.
**Estimated from Figure 7, of the study by Day (19, p. 195).

the words occurring as the various parts of speech, regarded possessive pronouns and nouns in apposition as adjectives, whereas she considered the former as pronouns and the latter as nouns. She also indicated that according to her authority, the *Century Dictionary*, he classified certain words as conjunctions which she regarded as different parts of speech. Boyd, therefore, had more adjectives and conjunctions than did Nice, and the latter recorded more pronouns. As shown in Table 20, Boyd classified as adjectives 17.4 per cent at 36 months while Nice recorded 9.3 per cent at that age. At the same age level the respective percentages for pronouns were 15.8 and 22.3.

The valuable studies by McCarthy and Smith have relatively large numbers of subjects and, therefore, afford an important opportunity for comparison. Smith (78, p. 16) states concerning the authority for her work: "The classification of words used into parts of speech followed the usual grammatical rules, Webster's dictionary being the criterion in case of doubt." The writer is indebted to Dr. McCarthy for a statement which is quite similar to that just quoted, except for the fact that Webster's *Collegiate Dictionary* was consulted when confusion occurred.

The studies by Smith and McCarthy differ as to the proportions of adjectives, with which articles have been combined in Table 20. At the ages of 18, 30, and 42 months, with McCarthy's age groups combined to make them more comparable to Smith's, the proportions of adjectives range, respectively, from 10.0 to 15.2 and from 5.0 to 11.0. It is probable that the "miscellaneous" classification used by McCarthy and Day includes some words which were classified by other investigators as adjectives, and some which were regarded as adverbs. In the case of McCarthy's data these figures would add only one or two per cent, but for Day's material they would give from 8.5 to 12.5 per cent to be added to other classifications.

Comparisons of adverbs, using the ages previously indicated, result in differences of 5.1 to 10.5 per cent. The larger proportions of nouns found in McCarthy's data, 44.3 per cent as compared with 28 per cent for the group of 18-29 months, may be partly the result of the experimental procedure. The use of pictures and toys, and the employing of questions to elicit responses if they were not spontaneous, may have stimulated enumeration in McCarthy's investigation. It will be noted that the most marked differences in the proportions of nouns occur at the lower age levels where naming of objects is particularly characteristic.

Since similar techniques in securing and classifying data were employed by McCarthy and Day, it appears that differences in the subjects or in the grammatical classification may account for the major part of the differences between the investigations. These are greatest for the adjectives, adverbs, interjections, and miscellaneous classifications. As previously stated, Day's subjects were 80 pairs of twins.

A study of Table 20 will indicate that the following facts may be stated concerning the several studies:

1. As age increased nouns decreased. The decrease was most rapid at the early age levels. Although 50 per cent of the words were nouns at 18 months (McCarthy), the proportion at 30 months was 25.8 per cent, and at 54 months it was only 19.3 per cent. The trend was distinctly toward the proportions of nouns used by adults. Boyd found for the novelists that approximately 15 or 16 per cent of the words were nouns. The investigations involving a carefully controlled procedure in a restricted setting (McCarthy and Day) had somewhat larger proportions of nouns than did other studies.

2. Especially in the studies with subjects of 18 and 24 months of age, there was clearly demonstrated a tendency for pronouns to occur in larger proportions in the higher age levels, the most rapid rise in the curve being in the early stages. This also represented a tendency toward the language of adults, although several of the studies indicated that the pronouns used by the older children exceeded the percentages used by the novelists. Differences in interpretation of pronouns because of divergent views of grammarians undoubtedly account for some of the variations in the proportions of pronouns.

3. Adjectives, including articles, showed an increase with age in most of the investigations. Again, the trend was toward the pattern of adult language, as indicated by the proportions of the parts of speech. The proportions of adjectives found in the several investigations differ greatly. Again, differences in opinion as to what constitutes an adjective may explain many of the divergencies in the findings.

4. Verbs tended to increase with age, the greatest increase occurring at the early months. For the preschool subjects of the upper

age groups the proportions of verbs in several studies were higher than those for the adult groups.

5. Adverbs showed no marked change with age in Day's and McCarthy's studies, but a decrease occurred in Boyd's, Nice's, and Smith's investigations. In Boyd's and Day's studies the subjects at almost all ages tended to resemble the adults in regard to the proportion of adverbs used.

6. The proportion of prepositions used by younger subjects was much smaller than that for older subjects in most of the studies. None of the preschool groups reached the adult stage in use of this part of speech.

7. The proportion of conjunctions increased with age, although progress as indicated by some investigations was irregular. None of the groups attained the adult level in regard to the proportion of conjunctions.

8. In the studies which were based upon large groups of subjects the proportion of interjections showed a sharp decrease with age. The trend was toward the percentage found in adult usage.

Several summarizing statements in regard to Table 20 should be made. It was not unusual for the proportions of verbs and pronouns to excel the proportion of nouns after the third year. In several studies, approximately half of the child's language was made up of verbs and pronouns at the upper preschool age levels. At these ages the verbal behavior of the subjects was characterized by action words (verbs) and personal symbols, inasmuch as most of the pronouns were classified as personal pronouns. This is in accordance with Markey's (63, p. 86) observation. It was characteristic of most of the studies that the proportions of the parts of speech changed most rapidly before the age of three, and after that age the rate of change tended to be slower. The "dearth of prepositions, conjunctions, articles, and auxiliary verbs" as emphasized by Boyd (8) was outstanding at the early age levels. The proportions of nouns and interjections decreased with age, but verbs, pronouns, adjectives, prepositions, and conjunctions usually showed an increase with age. The proportions of the parts of speech were characterized by rather large amounts of variation from investigation to investigation, probably due in part to the fact that different grammars were used as the bases of classification.

TABLE 21

MEAN PER CENT OF EACH PART OF SPEECH BY *CA*, GROUP, AND SEX (BASED ON THE TOTAL NUMBER OF COMPREHENSIBLE WORDS USED IN 144 HOURS,* OUTDOOR RECORDS PREDOMINATING)

CA	Group	Sex	Nouns	Pron.	Art.	Verbs	Adv.**	Conj.	Prep.	Adj.	Infin.**	Interj.
30	Reg.	M.	16.2	25.9	1.0	24.9	14.2	1.2	3.8	6.3	3.4	2.6
	Rel.	M.	20.2	25.8	0.5	27.5	10.3	0.5	2.9	6.1	1.9	3.5
	Reg.	F.	16.5	26.6	1.3	24.1	10.4	1.5	3.9	7.7	3.9	3.7
	Rel.	F.	18.6	27.2	1.6	27.1	7.0	1.0	3.7	7.5	2.8	3.3
	All	M.	18.2	25.9	0.8	26.2	12.3	0.9	3.4	6.2	2.7	3.1
	All	F.	17.6	26.9	1.5	25.6	8.7	1.3	3.8	7.6	3.4	3.6
	All		17.9	26.4	1.2	25.9	10.5	1.1	3.6	6.9	3.1	3.3
36	Reg.	M.	16.7	25.9	1.6	23.3	13.0	1.4	4.0	7.1	3.6	3.0
	Rel.	M.	20.4	26.2	1.4	26.0	9.1	0.8	3.2	7.4	2.3	2.5
	Reg.	F.	14.9	24.7	2.5	26.4	9.0	1.7	4.9	8.0	4.7	3.2
	Rel.	F.	17.4	26.0	2.1	28.2	8.6	0.9	3.1	7.2	3.1	3.6
	All	M.	18.6	26.1	1.5	24.7	11.1	1.1	3.6	7.3	3.0	2.8
	All	F.	16.2	25.4	2.3	27.3	8.8	1.3	4.0	7.6	3.9	3.4
	All		17.4	25.8	1.9	26.0	10.0	1.2	3.8	7.5	3.5	3.1
42	Reg.	M.	14.8	25.4	1.7	25.5	13.0	2.3	4.2	7.0	3.1	3.0
	Rel.	M.	18.3	27.3	1.8	25.0	9.3	1.6	3.6	7.9	2.2	3.0
	Reg.	F.	16.6	24.1	2.4	25.7	10.5	2.7	4.7	6.5	3.9	2.5
	Rel.	F.	15.7	26.6	2.0	27.7	8.0	1.4	3.7	7.5	3.1	3.6
	All	M.	16.6	26.4	1.8	25.3	11.2	2.0	3.9	7.5	2.7	3.0
	All	F.	16.2	25.4	2.2	26.7	9.3	2.1	4.2	7.0	3.5	3.1
	All		16.4	25.9	2.0	26.0	10.3	2.1	4.1	7.3	3.1	3.1
48	Reg.	M.	14.1	23.5	2.4	25.9	12.1	2.0	4.6	6.9	3.7	3.8
	Rel.	M.	19.5	22.8	2.5	26.9	10.2	1.4	4.1	7.6	2.6	2.4
	Reg.	F.	11.5	29.0	2.2	25.8	8.6	2.8	5.9	7.2	3.4	3.6
	Rel.	F.	12.2	29.5	2.8	27.4	9.3	1.5	4.8	6.5	3.0	3.0
	All	M.	16.8	23.2	2.5	26.4	11.2	1.7	4.4	7.3	3.2	3.1
	All	F.	11.9	29.3	2.5	26.6	9.0	2.1	5.4	6.9	3.2	3.3
	All		14.4	26.3	2.5	26.5	10.1	1.9	4.9	7.1	3.2	3.2

TABLE 21 (*continued*)

CA	Group	Sex	Nouns	Pron.	Art.	Verbs	Adv.**	Conj.	Prep.	Adj.	Infin.**	Interj.
54	Reg.	M.	12.3	26.5	2.3	24.5	12.7	1.8	4.7	6.6	4.1	3.9
	Rel.	M.	17.5	24.1	2.3	26.7	12.4	0.7	3.8	7.6	2.7	1.2
	Reg.	F.	11.4	25.8	3.5	27.6	10.8	2.5	5.2	7.3	3.2	2.7
	Rel.	F.	12.3	25.6	3.3	25.6	13.9	2.0	4.9	7.4	3.7	1.3
	All	M.	14.9	25.3	2.3	25.6	12.6	1.3	4.3	7.1	3.4	2.6
	All	F.	11.9	25.7	3.4	26.6	12.4	2.3	5.1	7.4	3.5	2.0
	All		13.4	25.5	2.9	26.1	12.5	1.8	4.7	7.3	3.5	2.3
60	Reg.	M.	15.5	25.4	3.9	24.2	11.2	2.4	4.2	6.4	3.9	2.2
	Rel.	M.	21.0	22.6	3.0	24.8	11.7	1.8	4.0	5.9	3.4	1.4
	Reg.	F.	13.6	26.4	3.7	23.5	12.4	2.4	4.9	6.9	4.2	1.7
	Rel.	F.	16.1	25.2	3.2	23.1	11.8	2.2	5.3	7.2	3.7	1.8
	All	M.	18.3	24.0	3.5	24.5	11.5	2.1	4.1	6.2	3.7	1.8
	All	F.	14.9	25.8	3.5	23.3	12.1	2.3	5.1	7.1	4.0	1.8
	All		16.6	24.9	3.5	23.9	11.8	2.2	4.6	6.7	3.9	1.8
30 to 60	Reg.	M.	14.8	25.4	1.8	24.8	13.0	1.7	4.3	6.8	3.6	3.3
	Rel.	M.	19.2	25.2	1.7	26.4	10.3	1.0	3.5	7.3	2.3	2.5
	Reg.	F.	14.2	26.0	2.4	25.9	9.9	2.2	4.9	7.3	3.8	3.1
	Rel.	F.	15.2	27.0	2.4	27.2	9.4	1.3	4.0	7.2	3.1	3.0
	All	M.	17.0	25.3	1.8	25.6	11.7	1.4	3.9	7.1	3.0	2.9
	All	F.	14.7	26.5	2.4	26.6	9.7	1.8	4.5	7.3	3.5	3.1
	All		15.9	25.9	2.1	26.1	10.7	1.6	4.2	7.2	3.3	3.0
30 to 65	Reg.	M.	14.9	25.4	2.2	24.7	12.7	1.9	4.3	6.7	3.6	3.1
	Rel.	M.	19.5	24.8	1.9	26.2	10.5	1.1	3.6	7.1	2.5	2.3
	Reg.	F.	14.1	26.1	2.6	25.5	10.3	2.3	4.9	7.3	3.9	2.9
	Rel.	F.	15.4	26.7	2.3	26.5	9.8	1.5	4.3	7.2	3.2	2.8
	All	M.	17.2	25.1	2.1	25.5	11.6	1.5	4.0	6.9	3.1	2.7
	All	F.	14.8	26.4	2.6	26.0	10.1	1.9	4.6	7.3	3.6	2.9
	All		16.0	25.8	2.4	25.8	10.9	1.7	4.3	7.1	3.4	2.8

*Approximately 50 per cent of the records were made *Outdoors*.

**The *Infinitive* group also includes the participial and gerundial constructions. The term *Adverb* includes pronominal adverbs.

B. ANALYSIS OF DATA

1. *Proportion of Each Part of Speech*

As stated in a previous chapter, records covering 80 minutes of verbal behavior per subject were obtained in addition to the 280 minutes per child composing the uniform records. Table 21 is based upon the 360 minutes of observation for each of the 74 subjects, a total of 444 hours. The uniform and miscellaneous records are combined for this portion of the investigation, and the material differs from the discussion of the parts of speech in other parts of the study in the following respects. The four settings do not have equal weight in this treatment of the data, as they do where only the uniform records are used. Since the Outdoor records are the most numerous, they have the predominant influence in this part of the investigation. In Table 21, the age groups are equally represented in the group means, although the groups are not composed of the same numbers of subjects. In this portion of the study, percentages are the basis of discussion, whereas the discussion of the records which will accompany the tables showing the critical ratios will be based upon the actual incidence of the parts of speech in verbal behavior.

Since the group at 60 months apparently was not representative, and because of the small number composing it, the means are given for ages 30 to 60 months as well as for ages 30 to 65 months. Discussion of Table 21 will involve certain comparisons with Table 20 in which were given the results of several other investigations. In the present study, the following points should be noted:

a. Nouns. In accordance with the findings of other studies, there is a decrease in the proportion of nouns with an increase in age. The mean at 30 months is 17.9 per cent, and at 54 months it is 13.4 per cent. The group at 60 months is small, and the rise to 16.6 per cent probably should not be regarded as significant. The percentage of nouns is smaller than that found by McCarthy and Day (Table 20), but it will be recalled that they followed a procedure which involved a setting similar to the Picture situation of the present study, whereas the largest proportion of the records for Table 21 was made in the Outdoor situation. The proportion of nouns is somewhat similar to the percentages found by Boyd, Nice, and Smith (Table 20). The two former used many settings, and the last-named used both outdoor play with children and indoor play involving contact with adults, in securing her records.

In the present study, Relief subjects have a somewhat larger proportion of nouns than Regular subjects. McCarthy (57, p. 123) states concerning her results that "nouns are a higher percentage of the total number of words used by the children of this study who belong to the lower occupational group." McCarthy regards a smaller proportion of nouns in the older subjects as an indication of "a higher stage of linguistic development." In accordance with this statement, girls in the present study are slightly superior to boys. Observation reveals, however, that Regular girls, Regular boys, and Relief girls have similar records, and all are superior to Relief boys. It is, therefore, Relief boys who are largely responsible for the greater proportion of nouns for Relief subjects and for boys as a group. For the combined groups, nouns constitute 16 per cent of the comprehensible words used in 444 hours.

b. Pronouns. Inasmuch as the present study includes no subjects less than 30 months of age, the increase in pronouns which is observed at the early age levels in Table 20 is not characteristic of Table 21. In the present investigation, pronouns maintain approximately the same proportion at all ages from 30 to 65 months, the mean for all ages being 25.8 per cent. The percentages are larger than those in some studies, the results more nearly resembling those of Nice and Smith (Table 20) than those given by other investigators. It will be recalled that the present study is based upon Jespersen's (45) viewpoint, and he classifies as pronouns certain words which are regarded as other parts of speech by some authorities. This point has been discussed in a previous chapter and is an important fact for consideration in comparisons of investigations. Probably some of the differences in the findings are explicable on the basis of the divergent views concerning grammar. That the differences are not greater may be due to the predominance of such pronouns as *I, me, you, he, she,* and *it,* concerning the classification of which there is not much dispute.

In the present study of the proportions of pronouns used, the differences between the sexes and between the socio-economic groups are very small.

c. Articles. Articles increase from 1.2 per cent at 30 months of age to 3.5 per cent at 60 months, the mean for all ages being 2.4 per cent. Girls are superior to boys in the use of articles, and Regular subjects somewhat surpass Relief subjects. Relief boys use the fewest articles when compared with other groups, the inferiority being greatest at 30 and 60 months. The superiority of girls in the proportion of articles used is especially noticeable at the earlier ages and at 54 months.

d. Verbs. Since no cases less than 30 months of age are included, an increase in verbs as observed in some of the investigations reported in Table 20 is not noted in the present study. The proportions are quite similar to those obtained by Smith, and they do not differ greatly from those noted in the other studies. When the percentage of infinitives is added to the proportion of verbs the results most nearly resemble those of Boyd, Kirkpatrick, and Nice. The percentage for the combined group is 29.2 per cent. Without the addition of infinitives, verbs compose 25.8 per cent of the comprehensible verbal behavior of the subjects in the present study.

e. Adverbs. The records of adverbs change little with age until 54 months when there is an increase. Group differences are largest prior to that age, and at most ages Regular boys show noticeable superiority in the proportion of adverbs. The mean for all ages is 10.9 per cent.

The results of the present study more nearly resemble those of Boyd and Day (Table 20) than those of other investigators. It should be observed that McCarthy and Day classified as *miscellaneous* certain words which some investigators may have regarded as adverbs or other parts of speech.

f. Conjunctions. Conjunctions increase from 1.1 per cent at 30 months to 2.2 per cent at 60 months. The mean for all ages is 1.7 per cent. Regular subjects are superior to Relief cases, and girls surpass boys.

The present investigation shows a smaller proportion of conjunctions than indicated by McCarthy or Boyd (Table 20), and fewer than other experimenters indicated at the upper age levels, with the exception of Nice's findings. Differences in grammatical classification probably account for some of the variations reported by experimenters.

g. Prepositions. Prepositions increase from 3.6 per cent at 30 months to 4.6 per cent at 60 months. Regular subjects surpass Relief cases and girls surpass boys in the proportion of prepositions, the differences varying at the several age levels. Most of the other investigators report a somewhat larger percentage of prepositions than is shown in the present study.

h. Adjectives. In the present study adjectives do not show a consistent increase with age after a slight rise at 36 months. They compose 7.1 per cent of the comprehensible words and if articles are added, as was the procedure in the investigations reported in Table 20, the increase with age is more consistent and the proportion is 9.5 per cent. The percentages are noticeably smaller than those obtained by Boyd and McCarthy. The proportions more closely resemble those of Day's study, and the records at the earlier ages from Nice and Smith. It should be recalled that in the present study certain words are classified as pronouns which other investigators may have regarded as adjectives. In a previous section may be found a discussion of this problem. Girls have a slightly larger proportion of adjectives than boys.

i. Infinitives. Infinitives compose 3.4 per cent of the speech of the subjects in the present investigation. No consistent change with age is observable, but a comparison of the groups at 30 and 60 months shows an increase from 3.1 per cent to 3.9 per cent. Regular subjects have a larger proportion of infinitives in their records than Relief cases at all ages. Girls surpass boys, with the greatest superiority appearing at the earlier ages.

j. Interjections. Interjections decrease from 3.3 per cent at 30 months to 1.8 per cent at 60 months. They constitute 2.8 per cent of the comprehensible speech of the subjects. The proportions more nearly resemble the findings, except at the earliest ages, of Day and Smith (Table 20) than those of other investigators. Regular subjects have a larger percentage than Relief cases, and girls a slightly larger percentage than boys.

2. *Incidence of Parts of Speech*

The present investigation includes only subjects aged 30 months or more and, therefore, the moot question of sentence-words is not as important for this study as for one which involves younger children.

It will be recalled that Jespersen's (45) statements concerning the parts of speech constitute the guiding principles in the classification used in the present study.

The incidence of the parts of speech is closely related to the number of comprehensible words spoken. In a previous chapter the change in vocal behavior concurrent with change in setting was discussed. It is assumed that those factors which affected the total number of words in the several settings also affect the individual parts of speech, although perhaps in different degrees in different situations.

Self-comparisons in the four different situations are presented for the several parts of speech, or groups of parts of speech, to show possible changes concurrent with change in setting. Likewise, group comparisons are made on the basis of the specific settings and, finally, on the basis of the combined situations.

The data in the succeeding discussions of this chapter are based upon the uniform records, and in these records the four settings have equal weight, as previously described.

Statements concerning variations in the parts of speech refer only to changes in the frequencies and not to differences in form.

Correlations which are involved in the computations of certain critical ratios may be found in Appendix *B*. All correlations have been calculated by the Pearson product-moment method.

a. Nouns. The reliability coefficients for the several parts of speech may be found in Table 22. The correlations for nouns range from .83 to .90 at the various ages.

TABLE 22

RELIABILITY COEFFICIENTS* OF DESIGNATED PARTS OF SPEECH IN ODD-EVEN-NUMBERED PERIODS FOR AGE GROUPS (BASED ON OUTDOOR-INDOOR UNIFORM RECORDS)

CA	Nouns	Pronouns	Verbs	Particles*	Adj.-Inf.*	*PE* Range
30	.83	.86	.82	.84	.81	.022-.017
36	.85	.84	.86	.87	.89	.022-.016
42	.84	.87	.88	.89	.86	.020-.014
48	.87	.83	.85	.84	.82	.023-.017
54	.90	.88	.86	.86	.88	.015-.011
60	.85	.82	.84	.82	.84	.027-.023

*Pearson product-moment correlations are given. The Particles group includes adverbs, conjunctions, and prepositions. The Adjectives-Infinitives group also includes participial and gerundial constructions, and interjections.

In Tables 23 and 24 are given the central tendencies and deviations for the parts of speech, according to groups and settings. The mean incidence of nouns for Regular boys, Relief boys, Regular girls and Relief girls is as follows: 9.23, 9.33, 11.41, and 10.22.

Correlations between the several combinations of settings for the

TABLE 23

CENTRAL TENDENCIES AND DEVIATIONS FOR PARTS OF SPEECH* PER 10-MINUTE PERIOD, AS USED PER INDIVIDUAL BY REGULAR BOYS AND RELIEF BOYS IN THE SPECIFIED SETTINGS

		Regular boys			Relief boys		
Part of Speech	Setting	Mean	*SD*	SD_m	Mean	*SD*	SD_m
Nouns	Out.	9.08	3.99	.45	9.43	4.19	.47
	In.	8.83	3.74	.42	8.94	3.63	.41
	Din.	7.90	3.15	.35	8.31	3.60	.40
	Pict.	11.11	4.75	.53	10.63	3.61	.40
Pronouns	Out.	12.38	4.55	.51	9.46	3.49	.39
	In.	10.71	4.45	.50	9.16	3.33	.37
	Din.	8.85	3.18	.36	8.00	3.42	.38
	Pict.	12.19	4.44	.50	9.58	3.45	.39
Verbs	Out.	11.41	4.79	.54	8.53	3.53	.39
	In.	9.06	3.95	.44	8.88	2.97	.33
	Din.	8.90	3.70	.41	7.80	3.14	.35
	Pict.	15.26	6.36	.71	9.35	3.71	.41
Particles	Out.	8.89	4.24	.47	7.78	3.05	.34
	In.	7.64	3.24	.36	6.31	2.77	.31
	Din.	6.91	3.06	.34	5.76	2.88	.32
	Pict.	10.56	4.67	.52	6.75	3.08	.34
Adj.-Inf.	Out.	7.46	3.18	.36	6.23	3.41	.38
	In.	6.61	3.07	.34	5.98	2.96	.33
	Din.	6.65	2.88	.32	6.74	3.11	.35
	Pict.	7.48	3.08	.34	6.48	2.53	.28

*The Pronoun group also includes articles. The Particles group is composed of adverbs, conjunctions and prepositions. The Adjective-Infinitive group also includes the participial and gerundial constructions, and interjections.

various groups may be found in Appendix *B*. The 24 correlations range from .73 to .89, indicating a marked degree of consistency in the occurrence of nouns from setting to setting. The subjects who use large numbers of this part of speech in one situation tend to maintain that characteristic in the other settings, whereas those whose records show a relative paucity of nouns in one situation tend to have similar records in the other environments.

In Table 25 are given the "Critical Ratios of the Differences between Settings," for nouns as used by the several groups. The self-comparisons of Regular boys indicate that the incidence of nouns may differ markedly with the setting, since five of the six differences

TABLE 24

Central Tendencies and Deviations for Parts of Speech* Per 10-Minute Period, as Used Per Individual by Regular Girls and Relief Girls in the Specified Settings

		Regular girls			Relief girls		
Part of Speech	Setting	Mean	*SD*	SD_m	Mean	*SD*	SD_m
Nouns	Out.	12.96	4.37	.53	8.69	3.49	.42
	In.	10.72	3.82	.46	9.51	3.18	.39
	Din.	9.60	3.12	.38	8.81	3.77	.46
	Pict.	12.35	4.33	.53	13.87	4.61	.56
Pronouns	Out.	19.93	7.85	.95	12.50	4.73	.57
	In.	18.38	6.74	.82	11.66	4.26	.52
	Din.	12.97	4.49	.54	11.13	3.78	.46
	Pict.	17.75	6.96	.84	12.59	4.77	.58
Verbs	Out.	17.09	5.70	.69	11.13	4.93	.60
	In.	13.13	4.33	.53	9.96	3.81	.46
	Din.	9.15	3.05	.37	9.07	4.11	.50
	Pict.	12.25	5.09	.62	11.59	4.14	.50
Particles	Out.	12.82	4.41	.53	8.74	3.10	.38
	In.	9.65	3.71	.45	7.54	3.31	.40
	Din.	8.21	2.76	.33	6.76	3.05	.37
	Pict.	9.75	3.18	.39	7.41	3.12	.38
Adj.-Inf.	Out.	12.22	4.40	.53	7.32	2.60	.32
	In.	7.88	2.95	.36	6.84	2.21	.27
	Din.	7.51	3.26	.40	6.74	2.60	.32
	Pict.	8.97	3.54	.43	7.22	2.69	.33

*The Pronoun group also includes articles. The Particles group is composed of adverbs, conjunctions and prepositions. The Adjective-Infinitive group also includes the participial and gerundial constructions, and interjections.

TABLE 25

Critical Ratios of the Differences Between Settings in the Number of Nouns Used Per 10-Minute Period by the Group Specified (Correlated Observations)

Setting	In.	Pict.	Din.	Setting	In.	Pict.	Din.
Regular boys				Relief boys			
Pict.	9.12			Pict.	6.04		
Din.	4.43	10.70		Din.	2.43	8.00	
Out.	1.09	7.25	4.54	Out.	2.33	4.29	4.00
Regular girls				Relief girls			
Pict.	5.09			Pict.	13.63		
Din.	4.00	8.33		Din.	2.41	15.81	
Out.	9.33	1.65	10.84	Out.	3.28	15.24	0.44

satisfy the common criterion for statistical reliability. The situations with the larger means will be mentioned first in the comparisons which follow: Outdoor-Dinner; Indoor-Dinner; Picture-Dinner; Picture-Outdoor; and Picture-Indoor.

The fact that the critical ratio is only 1.09 as a result of the comparison of nouns in the Outdoor-Indoor settings should be noted, since the comparison of the amount of comprehensible verbal behavior in those situations, as shown in a previous chapter, resulted in a significant difference. In both instances the largest incidence is to be found in the Picture situation.

If the observations are treated as if uncorrelated, only the critical ratios for two comparisons are indicative of statistically reliable differences: Picture-Dinner and Picture-Indoor. However, the critical ratios for the Picture-Outdoor and Outdoor-Dinner comparisons indicate approximately 99.8 and 98 chances in 100 that the true differences are greater than zero.

Resulting from the self-comparisons of nouns as used in the various settings by Relief boys (Table 25) are four differences which are so large, as measured by their standard errors, that they may be regarded as statistically significant. These four comparisons are as follows: Picture-Outdoor; Picture-Indoor; Picture-Dinner; and Outdoor-Dinner. The critical ratios of 2.42 and 2.33 for the Indoor-Dinner and Outdoor-Indoor comparisons indicate approximately 99.2 and 99.0 chances in 100 that true differences exist. It will be recalled that the self-comparisons of Relief boys as to number of words spoken in the various settings (Table 26) resulted in five highly reliable differences, and the sixth critical ratio indicated virtual significance.

If the observations are treated as if uncorrelated, only the difference between the Picture-Dinner settings is completely reliable statistically. However, the critical ratios for the comparisons of the Outdoor-Dinner, Picture-Outdoor, and Picture-Indoor settings indicate close approximations to statistical significance since there are, respectively, 96, 97.5, and 99.9 chances in 100 that a true difference exists.

Regular girls differ from setting to setting as to the number of nouns used, and five of the six critical ratios are indicative of highly significant differences, as follows: Outdoor-Indoor; Outdoor-Dinner; Indoor-Dinner; Picture-Indoor; and Picture-Dinner. It should be noted that the critical ratio resulting from the comparison of the

TABLE 26

CRITICAL RATIOS OF THE DIFFERENCES BETWEEN COMPARED GROUPS IN MEAN NUMBER OF NOUNS PER 10-MINUTE PERIOD IN THE SETTINGS SPECIFIED

Setting	Out.	In.	Din.	Pict.	All	Setting	Out.	In.	Din.	Pict.	All
	Reg. boys—Rel. boys						Reg. girls—Rel. girls				
Out.	0.54					Out.	6.28				
In		0.19				In.		2.02			
Din.			0.77			Din.			1.32		
Pict.				0.73		Pict.				1.97	
All					0.18	All					1.98
	Reg. boys—Reg. girls						Rel. boys—Rel. girls				
Out.	5.54					Out.	1.17				
In.		3.05				In.		1.00			
Din.			3.27			Din.			0.82		
Pict.				1.65		Pict.				4.70	
All					3.70	All					1.56
	Reg. boys—Rel. girls						Rel. boys—Reg. girls				
Out.	0.63					Out.	4.97				
In.		1.19				In.		2.87			
Din.			1.57			Din.			2.35		
Pict.				3.58		Pict.				2.61	
All					1.68	All					3.59

number of words spoken (Table 26) in the Outdoor-Picture settings likewise indicated a highly reliable difference. However, the comparison of nouns yields a critical ratio of 1.65, which gives approximately 95 chances in 100 that a true difference exists. When the Picture-Indoor settings are compared as to amount of comprehensible verbal behavior (Table 26) the critical ratio is 1.92, and indicative of approximately 97 chances in 100 that the true difference is greater than zero. When comparison is made of noun incidence in these same settings the critical ratio of 5.09 reveals a highly reliable difference. It is possible that nouns were elicited to such a degree in the Picture setting as to produce noticeable changes in the relationships existing from comparisons of the amount of comprehensible verbal behavior.

Treatment of the observations as if uncorrelated results in the following statistically significant differences between settings: Picture-Dinner; Outdoor-Dinner; and Outdoor-Indoor. Critical ratios of 1.87 and 2.33 for the Indoor-Dinner and Picture-Indoor comparisons indicate approximately 97 and 99 chances in 100 that true differences exist.

In the self-comparisons of Relief girls in the several settings as to noun incidence, four of the six critical ratios satisfy the usual criterion for statistical reliability, as follows: Picture-Outdoor; Picture-Indoor; Picture-Dinner; and Indoor-Outdoor. The critical ratio from the comparison of the Indoor-Dinner settings indicates approximately 99.2 chances in 100 of a real difference. It should be noted that the comparison of the number of words spoken (Table 26) in the Outdoor-Dinner settings resulted in a highly reliable difference, although the noun frequencies in these situations do not differ greatly.

If the observations are regarded as if uncorrelated, only the following comparisons yield statistically reliable differences: Picture-Dinner; Picture-Indoor; and Picture-Outdoor.

The groups tended to respond to the Picture situation with an observable amount of enumeration. Enumeration in such a situation is characteristic of this age, as noted by Binet, Terman (83), and others. The tendency probably was accentuated by the use of questions to elicit speech, if the subjects did not talk spontaneously. These factors may account in part for the fact that the noun incidence for three of the groups is greatest in the Picture setting, although the amount of verbal behavior likewise is greatest in this situation in the case of these groups, as reported in an earlier chapter.

Regular girls differ from the other groups in having the greatest amount of speech and the largest noun incidence in the Outdoor situation. Regular girls gave a somewhat higher type of response in the Picture setting than did the other subjects. They often used some description, and this is considered to be a more mature type of reaction than is mere enumeration. The group was characterized by more spontaneous remarks than were other subjects and, therefore, it was not necessary to elicit conversation to the extent that this was required for the other groups.

Table 26 gives the critical ratios of the differences between groups as regards the incidence of nouns in the four settings. Regular boys and Relief boys do not differ significantly in the specific or combined settings, the largest critical ratio being 0.77.

Regular girls surpass all other groups in noun incidence, as would be expected from their superiority in amount of verbal behavior. Regular girls and Regular boys differ markedly in all settings, and comparisons of noun incidence in the Outdoor, Indoor, Dinner, and combined settings result in significant differences. These findings correspond to those obtained when the number of words spoken constituted the basis of comparison, as previously discussed.

Regular girls and Relief girls differ markedly in the Indoor, Picture, and combined settings, and the critical ratios indicate approximately 98 chances in 100 that real differences exist. Only the comparison in the Outdoor setting yields a critical ratio which satisfies the criterion for significance. When the two groups were compared *on the basis of the number of words spoken, the critical ratios ranged* from 1.80 to 7.27, and three of the five were indicative of reliable differences (Table 17).

Regular girls surpass Relief boys in noun incidence in all settings, and the critical ratios indicate approximately 99 or more chances in 100 that true differences exist. In the Outdoor and combined settings the differences apparently are statistically reliable. It will be recalled that all critical ratios satisfied the criterion for significance when the groups were compared on the basis of the amount of verbal behavior (Table 17).

As was the case when Relief boys and Relief girls were compared on the basis of amount of verbal behavior, only the critical ratio in the Picture setting is indicative of a completely significant difference between groups in the incidence of nouns. However, the critical ratios obtained from comparisons of the number of words spoken by the two groups are somewhat larger than those from a similar comparison of nouns in the Outdoor, Indoor, Dinner, and combined settings.

Relief girls and Regular boys, as would be expected from comparisons of amount of verbal behavior, do not differ significantly in noun incidence in the Outdoor, Indoor, Dinner, and combined settings. Although the critical ratio obtained from comparing the amount of language used in the Picture setting is only 1.14 (Table 17), there is a statistically reliable difference between the noun usage of the two groups in this situation.

Comparisons between the sexes with the socio-economic factor constant (Regular girls-Regular boys, and Relief girls-Relief boys) on the basis of the combined settings yield, respectively, a significant and an insignificant difference.

Comparisons between socio-economic groups with the sex factor constant (Regular boys-Relief boys, and Regular girls-Relief girls) yield insignificant differences on the basis of the combined settings.

As previously stated, none of the comparisons of Regular boys-Relief boys in the specific settings results in reliable differences, and the largest critical ratio is 0.77. However, the critical ratios obtained from comparisons of Regular girls-Relief girls in specific settings range from 1.97 to 6.28. These statements are of especial interest since Drever (22), McCarthy (57), and others have discussed the relationship of socio-economic status to noun incidence, as noted earlier in this chapter.

Regular girls surpass all other groups in noun usage, and comparisons with Regular boys and Relief boys result in reliable differences. Relief girls rank second in noun incidence, but they do not significantly surpass Regular boys and Relief boys. The relative standing of the groups in regard to noun incidence (Table 26) is somewhat similar to that observed in comparing the amounts of comprehensible verbal behavior (Table 17). Most of the discrepancies which occur are associated with relatively large noun records on the part of Relief boys, and the apparent influence of the Picture setting upon the noun incidence of the groups.

In the study of noun incidence it should be observed that more significant differences result from self-comparisons of the same subjects in different settings than from comparisons of different groups in the same setting. The table of self-comparisons, Table 25, indicates 18 significant differences out of a possible 24, when correlated observations are used. The table of group comparisons, Table 26, indicates only seven significant differences between groups in the same setting, out of a possible 24. However, if the observations for the self-comparisons are treated as if uncorrelated, only nine completely reliable differences result from the self-comparisons in the several settings.

b. Pronouns. The classification of pronouns discussed in this section includes articles, as previously explained. The reliability coefficients for pronouns are given in Table 22, and they range from .82 to .88. These correlations are indicative of satisfactory reliability at all ages.

In Tables 23 and 24 are given the central tendencies and deviations for pronouns, according to settings. The mean incidence of pronouns for Regular boys, Relief boys, Regular girls, and Relief girls is as follows: 11.03, 9.05, 17.26, and 11.97.

The correlations between the several combinations of settings for the various groups may be found in Appendix *B*. The 24 correlations range from .72 to .90, indicating that the subjects who employ large, moderate, or small numbers of pronouns in one situation tend to demonstrate like reactions in the other settings.

In Table 27 are given the "Critical Ratios of the Differences between Settings," for pronouns as used by the various groups. The self-comparisons of Regular boys indicate that the incidence of pronouns may differ from setting to setting. Five of the six differences satisfy the criterion for statistical reliability. The situation with the larger mean will be mentioned first in the comparisons which follow: Picture-Indoor; Picture-Dinner; Outdoor-Indoor; Outdoor-Dinner; and Indoor-Dinner.

TABLE 27

CRITICAL RATIOS OF THE DIFFERENCES BETWEEN SETTINGS IN THE NUMBER OF PRONOUNS USED PER 10-MINUTE PERIOD BY THE GROUP SPECIFIED (CORRELATED OBSERVATIONS)

Setting	In.	Pict.	Din.	Setting	In.	Pict.	Din.
	Regular boys				Relief boys		
Pict.	5.69			Pict.	1.75		
Din.	7.15	11.93		Din.	4.64	5.45	
Out.	6.19	0.70	11.77	Out.	1.36	0.57	5.62
	Regular girls				Relief girls		
Pict.	1.70			Pict.	3.10		
Din.	12.30	9.96		Din.	1.66	4.87	
Out.	3.60	4.64	12.43	Out.	2.55	0.26	4.15

The critical ratios are similar to those obtained for the total number of words spoken except that for the latter the difference between the Outdoor-Picture settings is statistically reliable (Table 16).

If the observations are treated as if uncorrelated, only the critical ratios for the following comparisons are indicative of statistically reliable differences: Outdoor-Dinner; Indoor-Dinner; and Picture-Dinner. However, the critical ratios for the Outdoor-Indoor, and Picture-Indoor comparisons indicate, respectively, 99 and 98 chances in 100 that the true differences are greater than zero.

The self-comparisons of Relief boys (Table 27) result in three statistically significant differences between settings, as follows: Picture-Dinner; Indoor-Dinner; and Outdoor-Dinner. The Outdoor-Indoor, and Picture-Indoor comparisons yield critical ratios which indicate, respectively, 91 and 96 chances in 100 that a real difference exists. A critical ratio of 0.57 is the result of the Picture-Outdoor comparison. It will be noted that this is much smaller than the critical ratio obtained from a comparison of the amount of verbal behavior (Table 16).

When the situations are regarded as uncorrelated all critical ratios fail to satisfy the criterion for statistical significance. However, those for the Indoor-Dinner, Outdoor-Dinner, and Picture-Dinner settings indicate, respectively, 98.6, 99.6, and 99.8 chances in 100 that true differences exist.

In the self-comparisons of Regular girls in the various settings five of the six critical ratios are indicative of highly significant differences in pronoun incidence, as follows: Outdoor-indoor; Outdoor-Dinner; Outdoor-Picture; Indoor-Dinner; and Picture-Dinner. The results correspond to those obtained from a similar comparison of the number of words spoken (Table 16).

If the observations are treated as if uncorrelated the number of completely reliable differences is reduced to three: Outdoor-Dinner; Indoor-Dinner; and Picture-Dinner.

The self-comparisons of Relief girls in the various settings result in three significant differences in pronoun incidence, as follows: Outdoor-Dinner; Picture-Dinner; and Picture-Indoor. The critical ratios from the comparisons of the Indoor-Dinner, and Outdoor-Indoor settings indicate approximately 95 and 99.5 chances in 100 that true differences exist. It will be recalled that comparisons of the amount of comprehensible verbal behavior in the several settings (Table 16) resulted in statistically reliable differences in all instances, but the critical ratio from comparing pronouns in the Picture-Outdoor settings is only 0.26.

When the situations are regarded as uncorrelated the critical ratios fail to satisfy the criterion for significance. However, those from the Outdoor-Dinner, and Picture-Dinner comparisons indicate approximately 97 and 98 chances in 100 that the true differences are greater than zero.

In Table 28 may be found the "Critical Ratios of the Differences between Compared Groups in Mean Number of Pronouns," in the specific and combined settings. Regular boys and Relief boys differ significantly in pronoun usage in the Outdoor, Picture, and combined settings. In addition, the critical ratios of 2.50 and 1.63 in the

TABLE 28

Critical Ratios of the Differences Between Compared Groups in Mean Number of Pronouns Per 10-Minute Period, in the Settings Specified

Setting	Out.	In.	Din.	Pict.	All	Setting	Out.	In.	Din.	Pict.	All
	Reg. boys—Rel. boys						Reg. girls—Rel. girls				
Out.	4.56					Out.	6.69				
In.		2.50				In.		6.93			
Din.			1.63			Din.			2.59		
Pict.				4.14		Pict.				5.06	
All					3.54	All					5.88
	Reg. boys—Reg. girls						Rel. boys—Rel. girls				
Out.	6.99					Out.	4.41				
In.		7.99				In.		3.91			
Din.			6.34			Din.			5.22		
Pict.				5.67		Pict.				4.30	
All					7.16	All					4.79
	Reg. boys—Rel. girls						Rel. boys—Reg. girls				
Out.	0.16					Out.	10.17				
In.		1.32				In.		10.24			
Din.			3.93			Din.			7.53		
Pict.				0.52		Pict.				8.78	
All					1.40	All					9.89

Indoor and Dinner settings indicate, respectively, 99.4 and 94 chances in 100 that true differences exist. It will be recalled that in comparisons of the amount of verbal behavior only the difference between groups in the Picture setting proved to have complete reliability, although the critical ratios for the Outdoor and combined settings indicated virtual significance (Table 17).

Regular girls significantly surpass all other groups in pronoun incidence on the basis of the combined settings. This is consistent with the fact that Regular girls significantly surpass all other groups in amount of verbal behavior. Also, in the specific settings, comparisons of Regular girls-Regular boys, and Regular girls-Relief boys result in reliable differences in pronoun usage. When Regular girls-Relief girls are compared all differences are statistically significant except that in the Dinner setting, and this critical ratio indicates approximately 99.5 chances in 100 that a true difference exists. It will be recalled that comparisons of Regular girls with the other three groups on the basis of amount of verbal behavior resulted in 12 differences, from a possible 15, which were completely reliable statistically (Table 17).

Regular boys and Relief girls have pronoun records of marked similarity, except in the Dinner setting. The close resemblance of these groups has been noted in a previous chapter in connection with the discussion of the number of words spoken.

When Relief boys and Relief girls are compared as to pronoun usage all critical ratios are indicative of reliable differences. This is consistent with the fact that comparisons of the amount of verbal behavior yield critical ratios indicating 98.7, or more, chances in 100 of true differences (Table 17).

Comparisons between relatively distinct socio-economic groups with the sex factor constant (Regular boys-Relief boys, and Regular girls-Relief girls) result in significant differences, on the basis of the combined settings.

Comparisons between the sexes with the socio-economic factor constant (Regular girls-Regular boys, and Relief girls-Relief boys) result in reliable differences, on the basis of the combined settings.

Group differences in pronoun incidence concomitant with difference in sex or in socio-economic status are reasonably consistent with the findings obtained from such comparisons of amount of verbal behavior. In the latter instances all differences are completely reliable or closely approach significance (Table 17).

Regular girls significantly surpass all other groups in pronoun incidence. Relief girls rank second, Regular boys having almost as good a record, and both surpass Relief boys with highly reliable differences, on the basis of the combined settings. The relative standing of the groups in regard to pronoun incidence is quite similar to that for the amount of comprehensible verbal behavior. However, the self-comparisons in certain specific settings do not show that pronoun usage is affected by change in setting in identically the same manner that total verbal behavior seems to be influenced. The most marked discrepancies are related to the fact that the pronoun records (Table 17) in the Outdoor and Picture settings are very much alike for Regular boys, Relief boys, and Relief girls, although comparisons of total comprehensible speech in these settings result in reliable or virtually significant differences (Table 16).

In the study of pronoun incidence it should be noted that almost the same number of significant differences result from self-comparisons of subjects in different types of setting as from comparisons of different groups in the same settings. The table of self-comparisons, Table 27, indicates 16 significant differences, from a possible 24, when correlated observations are used. The table of group comparisons, Table 28, indicates 18 significant differences, from a possible 24, when different groups are compared in the same setting. However, if the observations for the self-comparisons are treated as if uncorrelated, only six completely reliable differences result from the self-comparisons in the several settings. On this basis, more differences in pronoun incidence are concurrent with change in socio-economic status or in sex than with change in setting.

c. Verbs. The reliability coefficients for verbs are given in Table 22 and they range from .82 to .88 at the several age levels.

In Tables 23 and 24 are to be found the central tendencies and deviations for verbs, according to settings. The mean verb incidence for Regular boys, Relief boys, Regular girls, and Relief girls is as follows: 11.16, 8.64, 12.91, and 10.44.

The correlations between the compared settings for the various groups are given in Appendix *B*. The 24 correlations range from .73 to .93, and they indicate a marked degree of consistency in verb incidence from setting to setting.

The "Critical Ratios of the Differences between Settings in Mean Number of Verbs" as used by the various groups are presented in Table 29. The self-comparisons of Regular boys indicate that the verb frequencies differ greatly from setting to setting, as would be expected from the fact that significant differences in the amount of comprehensible speech are concurrent with variations in setting for this group (Table 16). Five of the six verb differences satisfy the criterion for statistical reliability: Outdoor-Indoor; Outdoor-Dinner; Picture-Indoor; Picture-Dinner; and Picture-Outdoor. The difference between verbs in the Indoor-Dinner settings is not reliable, although a significant difference exists when the amount of comprehensible speech in these settings is compared (Table 16).

If the observations are treated as if uncorrelated, the critical ratios are reduced in size but the same five differences remain statistically significant.

The self-comparisons of Relief boys (Table 29) result in four significant differences between verbs as used in the compared settings: Outdoor-Dinner; Picture-Dinner; Picture-Outdoor; and

TABLE 29

CRITICAL RATIOS OF THE DIFFERENCES BETWEEN SETTINGS IN THE NUMBER OF VERBS USED PER 10-MINUTE PERIOD BY THE GROUP SPECIFIED (CORRELATED OBSERVATIONS)

Setting	In.	Pict.	Din.	Setting	In.	Pict.	Din.
	Regular boys				Relief boys		
Pict.	16.32			Pict.	2.14		
Din.	0.70	14.45		Din.	5.40	6.74	
Out.	11.19	10.69	10.91	Out.	1.67	3.73	3.04
	Regular girls				Relief girls		
Pict.	2.26			Pict.	6.52		
Din.	11.71	7.21		Din.	3.30	7.64	
Out.	11.00	14.24	16.89	Out.	4.33	1.59	6.24

Indoor-Dinner. It should be noted that the differences between verbs in Indoor-Outdoor, and Picture-Indoor settings do not satisfy the criterion for statistical reliability, although they indicate approximately 96 and 98 chances in 100 that true differences exist. Comparisons of the amount of verbal behavior in these settings yield five reliable differences and the sixth closely approximates significance (Table 16).

No significant differences result when the observations are treated as if uncorrelated. However, the critical ratios from the comparisons of the Indoor-Dinner, and Picture-Dinner settings indicate approximately 99 and 99.8 chances in 100 of true differences.

In the self-comparisons of Regular girls in the various settings (Table 29) five of the six critical ratios are indicative of highly significant differences, as follows: Outdoor-Indoor; Outdoor-Picture; Outdoor-Dinner; Indoor-Dinner; and Picture-Dinner. These correspond to the results obtained from a comparison of the number of words spoken by the group (Table 16). As in the Picture-Indoor comparison of the number of words spoken the critical ratio for the verb comparison does not indicate a statistically significant difference but one which closely approaches the criterion for reliability.

Treatment of the observations as if uncorrelated yields the same reliable differences as previously given, although the critical ratios are reduced in size.

The verb incidence as recorded for Relief girls also shows changes which are concurrent with changes in setting, and the differences between the following settings prove to be statistically reliable:

Outdoor-Indoor; Outdoor-Dinner; Indoor-Dinner; Picture-Indoor; and Picture-Dinner. These findings correspond to those obtained from comparing the amount of comprehensible speech, except for the fact that all critical ratios are indicative of significant differences in the latter comparisons (Table 16).

The number of commands and requests given by the subjects in the Outdoor setting may account, at least in part, for the relatively large verb incidence in that situation. Such reactions apparently involve the use of verbs.

When the situations are regarded as uncorrelated only one of the differences, that of the Picture-Dinner situation, is significant. However, the critical ratios from the comparisons of the Outdoor-Dinner, and Picture-Indoor settings indicate approximately 99.6 and 99.2 chances in 100 that true differences exist.

Table 30 gives the critical ratios of the differences between the compared groups as regards the incidence of verbs in the various settings. In accordance with the tendencies noted in the comparisons

TABLE 30

CRITICAL RATIOS OF THE DIFFERENCES BETWEEN COMPARED GROUPS IN MEAN NUMBER OF VERBS PER 10-MINUTE PERIOD, IN THE SETTINGS SPECIFIED

Setting	Out.	In.	Din.	Pict.	All	Setting	Out.	In.	Din.	Pict.	All
	Reg. boys—Rel. boys						Reg. girls—Rel. girls				
Out.	4.30					Out.	6.55				
In.		0.33				In.		4.53			
Din.			2.04			*Din.*			0.13		
Pict.				7.21		Pict.				0.83	
All					4.13	All					3.48
	Reg. boys—Reg. girls						Rel. boys—Rel. girls				
Out.	6.45					Out.	3.61				
In.		5.90				In.		1.89			
Din.			0.45			Din.			2.08		
Pict.				3.20		Pict.				3.45	
All					2.46	All					3.00
	Reg. boys—Rel. girls						Rel. boys—Reg. girls				
Out.	0.35					Out.	10.84				
In.		1.41				In.		6.86			
Din.			0.26			Din.			2.65		
Pict.				4.22		Pict.				3.92	
All					1.03	All					6.89

of amounts of speech, Regular boys and Relief boys differ significantly in the Outdoor, Picture, and combined settings. In addition, the critical ratio of 2.04 in the Dinner situation indicates approximately 98 chances in 100 that the difference is a true one.

As would be expected from the fact that they have the greatest amount of verbal behavior, Regular girls markedly surpass all groups in verb incidence. When Regular girls are compared with Regular boys the critical ratios in the Outdoor, Indoor, and Picture settings indicate highly reliable differences, and that for the combined settings gives approximately 99.4 chances in 100 of the existence of a true difference. Regular girls and Regular boys have similar verb records in the Dinner situation, although they differ significantly in that setting in amount of comprehensible speech (Table 17). Comparisons of Regular and Relief girls yield reliable differences in the Outdoor, Indoor, and combined situations. As in comparisons of amount of comprehensible speech, marked differences between verbs result when Regular girls and Relief boys are compared. All critical ratios are indicative of highly reliable differences except in the Dinner situation, and that difference closely approaches statistical significance.

Regular boys and Relief girls have verb records of marked similarity, except in the Picture situation where Regular boys have a relatively large number of verbs and the critical ratio is 4.22. All group differences were relatively small when amount of speech was the basis of comparison (Table 17).

Relief boys and Relief girls differ significantly in verb usage in the Outdoor, Picture, and combined settings, and the critical ratios in the Indoor and Dinner settings indicate approximately 97 or 98 chances in 100 of the existence of true differences. It will be recalled that comparisons made on the basis of amount of verbal behavior yielded a reliable difference in the Picture setting and those resulting from all other comparisons closely approached significance (Table 17).

Comparisons between relatively distinct socio-economic groups with the sex factor constant (Regular boys-Relief boys, and Regular girls-Relief girls) result in significant differences, on the basis of the combined settings.

Comparisons between the sexes with the socio-economic factor constant (Regular girls-Regular boys, and Relief girls-Relief boys) result, respectively, in a virtually significant difference and a highly reliable difference, on the basis of the combined settings.

Group differences in verb incidence concomitant with differences in sex or in socio-economic status are relatively consistent with the findings obtained from such comparisons of amount of comprehensible speech (Table 17).

Regular girls markedly surpass all other groups in verb incidence, Regular boys rank second, Relief girls having almost as large a record, and both surpass Relief boys with reliable differences, on the basis of the combined settings. The relative standing of the groups in regard to verb incidence is somewhat similar to that for the amount of comprehensible verbal behavior. However, the self-comparisons in certain specific settings do not show that verb usage is affected by change in setting identically as the comprehensible vocal output seems to be influenced.

More significant differences in verb usage result from self-comparisons of subjects in different types of setting than from comparisons of different groups in the same settings. The table of self-comparisons, Table 29, indicates 19 significant differences, from a possible 24, when correlated observations are used. The table of group comparisons, Table 30, indicates 13 reliable differences, from a possible 24, when different groups are compared in the same settings. However, if the observations for self-comparisons are treated as if uncorrelated only 11 completely reliable differences result from the self-comparisons in the several settings. On this basis slightly more differences in verb incidence are concurrent with change in socio-economic status or in sex than with change in setting.

d. Particles. This term includes adverbs, conjunctions, and prepositions, as previously described. The reliability coefficients for particles are given in Table 22 and they range from .82 to .89. These correlations are indicative of satisfactory reliability at all ages.

In Tables 23 and 24 are given the central tendencies and deviations for particles according to settings. The mean incidence of particles for Regular boys, Relief boys, Regular girls, and Relief girls, is as follows: 8.50, 6.65, 10.11, and 7.61.

In Appendix *B* may be found the correlations between the compared settings for the various groups. The 24 correlations range from .75 to .89, indicating that, for the group as a whole, the subjects who employ large, moderate, or small numbers of particles in one situation tend to demonstrate similar reactions in the other settings.

In Table 31 are presented the "Critical Ratios of the Differences

TABLE 31

CRITICAL RATIOS OF THE DIFFERENCES BETWEEN SETTINGS IN THE NUMBER OF PARTICLES USED PER 10-MINUTE PERIOD BY THE GROUP SPECIFIED (CORRELATED OBSERVATIONS)

Setting	In.	Pict.	Din.	Setting	In.	Pict.	Din.
	Regular boys				Relief boys		
Pict.	11.23			Pict.	2.44		
Din.	3.48	12.17		Din.	3.06	5.50	
Out.	*4.46*	*6.19*	*6.39*	*Out.*	*6.39*	*4.90*	*9.18*
	Regular girls				Relief girls		
Pict.	0.36			Pict.	0.68		
Din.	4.80	6.70		Din.	3.71	3.25	
Out.	9.61	11.81	13.97	Out.	4.80	5.54	7.92

between Settings in the Number of Particles" as used by the various groups. The self-comparisons of Regular boys indicate that particles vary in number from setting to setting, as would be expected from the fact that significant differences in amount of comprehensible speech are concurrent with variations in setting for this group (Table 16). *All differences satisfy the criterion for statistical reliability* and the settings with the larger means will be mentioned first in the comparisons which follow: Outdoor-Indoor; Outdoor-Dinner; Picture-Outdoor; Indoor-Dinner; Picture-Indoor; and Picture-Dinner.

If the observations are treated as if uncorrelated the critical ratios are reduced in size and only the following indicate the presence of statistically reliable differences: Outdoor-Dinner; Picture-Indoor; and Picture-Dinner. However, the Outdoor-Indoor and Picture-Outdoor comparisons indicate, respectively, 98 and 99.2 chances in 100 that true differences exist.

The self comparisons of Relief boys (Table 31) result in one *virtually significant difference, that for the* Picture-Indoor settings, and five completely reliable differences, as follows: Outdoor-Indoor; Outdoor-Dinner; Outdoor-Picture; Indoor-Dinner; and Picture-Dinner. It will be recalled that self-comparisons of the amount of verbal behavior yield five completely reliable differences and the sixth closely approximates significance (Table 16).

Only two significant differences result when the observations are treated as if they are uncorrelated. However, the critical ratios from the Outdoor-Picture, and Picture-Dinner settings indicate, respectively, 98.3 and 98 chances in 100 of the probability of a real difference.

In the self-comparisons of Regular girls in the various settings (Table 31) five of the six critical ratios are indicative of highly significant differences, as follows: Outdoor-Indoor; Outdoor-Dinner; Outdoor-Picture; Indoor-Dinner; and Picture-Dinner. These findings are similar to the results obtained from self-comparisons of the number of words spoken by the group (Table 16).

If the observations are regarded as uncorrelated all differences just listed remain reliable except the difference resulting from the Indoor-Dinner comparison, and it closely approaches statistical significance.

The use of particles by Relief girls also shows changes which are concurrent with changes in setting, and the following differences between settings are reliable: Outdoor-Indoor; Outdoor-Dinner; Outdoor-Picture; Indoor-Dinner; and Picture-Dinner. These findings correspond to those obtained from self-comparisons of the amount of comprehensible speech, except for the fact that all critical ratios are indicative of significant differences in the latter comparisons. The Indoor-Picture records of particles are very similar, although there is a reliable difference between these settings in comprehensible verbal behavior (Table 16).

When the situations are regarded as uncorrelated only one of the differences, that from the Outdoor-Dinner comparison, is significant. However, the critical ratios from the comparisons of the Outdoor-Indoor, and Outdoor-Picture settings indicate, respectively, 98.6 and 99.4 chances in 100 that a true difference exists.

Table 32 gives the "Critical Ratios of the Differences between Compared Groups in Mean Number of Particles" in the various settings. When Regular boys and Relief boys are compared the differences obtained in the Picture and combined settings satisfy the criterion for statistical significance, and all other critical ratios indicate 97 or more chances in 100 that true differences exist. It will be recalled that comparisons of these groups as to amount of verbal behavior in the Outdoor, Picture, and combined settings result in differences which are reliable or which closely approach significance (Table 17).

In accordance with the fact that Regular girls have the greatest amount of verbal behavior, this group markedly excels all other groups in use of particles. When Regular girls are compared with Regular boys the critical ratios in the Outdoor and Indoor settings indicate reliable differences, and those in the Dinner and combined

TABLE 32

CRITICAL RATIOS OF THE DIFFERENCES BETWEEN COMPARED GROUPS IN MEAN NUMBER OF PARTICLES PER 10-MINUTE PERIOD, IN THE SETTINGS SPECIFIED

Setting	Out.	In.	Din.	Pict.	All	Setting	Out.	In.	Din.	Pict.	All
	Reg. boys—Rel. boys						Reg. girls—Rel. girls				
Out.	1.91					Out.	6.28				
In.		2.77				In.		3.52			
Din.			2.45			Din.			2.90		
Pict.				6.15		Pict.				4.25	
All					3.63	All					4.72
	Reg. boys—Reg. girls						Rel. boys—Rel. girls				
Out.	5.54					Out.	1.88				
In.		3.47				In.		2.41			
Din.			2.77			Din.			2.04		
Pict.				1.25		Pict.				1.29	
All					2.88	All					2.00
	Reg. boys—Rel. girls						Rel. boys—Reg. girls				
Out.	0.25					Out.	8.00				
In.		0.19				In.		6.07			
Din.			0.30			Din.			5.33		
Pict.				4.92		Pict.				5.77	
All					1.65	All					6.92

settings indicate virtual significance. As in comparisons of amount of verbal behavior the critical ratio in the Picture setting is smaller than in any others. Regular girls surpass Relief girls in all settings and differences in the Outdoor, Indoor, Picture, and combined situations are highly significant. A critical ratio of 2.90 in the Dinner setting indicates virtual significance. It will be recalled that differences from comparisons of amount of comprehensible speech as used by Regular girls-Relief girls are statistically reliable or closely approach significance (Table 17). As would be expected from the fact that reliable differences exist when the amount of verbal behavior is compared, all comparisons of particles as used by Regular girls and Relief boys result in highly significant differences.

Regular boys and Relief girls have records of particles which are markedly similar, except in the Picture situation where Regular boys have a relatively large incidence of particles and the critical ratio is 4.92. All group differences are unreliable when amount of speech is the basis of comparison between these groups (Table 17).

Relief boys and Relief girls do not differ significantly as to incidence of particles in any of the settings but all the critical ratios indicate 90 to 99.2 chances in 100 that true differences exist. It will be recalled that comparisons of these groups as to amount of verbal behavior yield a reliable difference in the Picture setting and those resulting from all other comparisons closely approach significance.

Comparisons between relatively distinct socio-economic groups with the sex factor constant (Regular boys-Relief boys, and Regular girls-Relief girls) result in reliable differences, on the basis of the combined settings.

Comparisons between the sexes with the socio-economic factor constant (Regular girls-Regular boys, and Relief girls-Relief boys) result in differences which indicate, respectively, 99.8 and 98 chances in 100 that a true difference exists, on the basis of the combined settings.

Group differences in incidence of particles concomitant with difference in sex or in socio-economic status are fairly consistent with the findings obtained from such comparisons of amount of comprehensible speech (Table 17).

Regular girls greatly surpass all other groups in use of particles. Regular boys rank second, and Relief girls third, both being markedly superior to Relief boys, on the basis of the combined settings. The relative standing of the groups in regard to the incidence of particles is somewhat similar to that for the amount of comprehensible speech. However, the self-comparisons in certain specific settings do not show that use of particles is affected by change in setting identically as comprehensible vocal output seems to be influenced.

More significant differences in use of particles result from self-comparisons of subjects in different types of setting than from comparisons of different groups in the same settings. The table of self-comparisons, Table 31, indicates 21 reliable differences, from a possible 24, when correlated observations are used. The table of group comparisons, Table 32, indicates 11 reliable differences, from a possible 24, when different groups are compared in the same settings. However, if the observations for self-comparisons are treated as if uncorrelated, only 10 completely reliable differences result from the self-comparisons in the several settings. On this basis approximately the same number of significant differences are concurrent with change in socio-economic status or in sex, as with variation in setting.

e. Adjectives-Infinitives. This term is used to include these parts of speech and also participial and gerundial constructions, and interjections. The reliability coefficients are given in Table 22, and they range from .81 to .89 at the several age levels.

In Appendix *B* are given the correlations between the compared settings for the various groups. The 24 correlations range from .76 to .90, and indicate a marked degree of consistency from setting to setting in the incidence of adjectives-infinitives.

In Tables 23 and 24 may be found the central tendencies and deviations for adjectives-infinitives, according to settings. The mean incidence for Regular boys, Relief boys, Regular girls, and Relief girls, is as follows: 7.05, 6.36, 9.15, and 7.03.

The "Critical Ratios of the Differences between Settings in the Number of Adjectives-Infinitives" as used by the various groups, are given in Table 33. The self-comparisons of Regular boys indicate that adjectives-infinitives vary in number from setting to setting, in accordance with the fact that significant differences in amount of verbal behavior are concurrent with all changes in situation (Table 16). The following critical ratios, obtained from self-comparisons of adjectives-infinitives, satisfy the criterion for statistical reliability: Outdoor-Indoor; Outdoor-Dinner; Picture-Indoor; and Picture-Dinner.

TABLE 33

CRITICAL RATIOS OF THE DIFFERENCES BETWEEN SETTINGS IN THE NUMBER OF ADJECTIVES-INFINITIVES USED PER 10-MINUTE PERIOD BY THE GROUP SPECIFIED (CORRELATED OBSERVATIONS)

Setting	*In.*	*Pict.*	*Din.*	*Setting*	*In.*	*Pict.*	*Din.*
	Regular boys				Relief boys		
Pict.	4.35			Pict.	2.50		
Din.	0.27	3.61		Din.	3.17	1.24	
Out.	4.05	0.09	4.26	Out.	1.25	1.04	2.13
	Regular girls				Relief girls		
Pict.	4.54			Pict.	1.81		
Din.	1.85	6.08		Din.	0.48	2.40	
Out.	14.97	11.61	16.24	Out.	2.67	0.48	2.64

If the observations are treated as if uncorrelated no differences are reliable, although all critical ratios from the comparisons listed above indicate approximately 96 chances in 100 of the possibility of true differences.

The self-comparisons of Relief boys result in one completely significant difference, that for the Dinner-Indoor settings, and the other critical ratios indicate from 85 to 99.4 chances in 100 of the probability of true differences. It will be recalled that self-comparisons of the amount of verbal behavior yield five reliable differences and the sixth closely approximates significance (Table 16).

When the observations are regarded as uncorrelated the largest critical ratio, 1.58, occurs in the Dinner-Indoor comparison, and it means that there are approximately 94 chances in 100 of a true difference.

In the self-comparisons of Regular girls in the various settings (Table 33) five of the differences prove to have statistical reliability. as follows: Outdoor-Indoor; Outdoor-Dinner; Outdoor-Picture; Picture-Indoor; and Picture-Dinner. The critical ratio from the Indoor-Dinner comparison indicates approximately 96.5 chances in 100 that the difference is a true one. The findings are somewhat similar to those obtained (Table 16) from self-comparisons of the number of words spoken by the group, and relatively large critical ratios in both instances are obtained from comparing the Outdoor setting with other settings.

If the observations are treated as if uncorrelated the differences between the Outdoor-Indoor, Outdoor-Dinner, and Outdoor-Picture settings attain the level of statistical reliability. The critical ratios from the Picture-Indoor, and Picture-Dinner comparisons indicate, respectively, 97.5 and 99.4 chances in 100 that a true difference exists.

Although Relief girls differ significantly from setting to setting in amount of verbal behavior (Table 16), corresponding self-comparisons of adjectives-infinitives result in no reliable differences. Critical ratios which indicate, however, approximately 96 to 99.7 chances in 100 of the existence of true differences result from comparisons of the following settings: Outdoor-Indoor; Outdoor-Dinner; Picture-Indoor; and Picture-Dinner.

When the observations are regarded as uncorrelated the critical ratios are reduced in size and the largest is 1.29, which indicates about 90 chances in 100 of a significant difference.

Table 34 gives the "Critical Ratios of the Differences between Compared Groups in Mean Number of Adjectives-Infinitives" in the various settings. When Regular boys and Relief boys are com-

TABLE 34

CRITICAL RATIOS OF THE DIFFERENCES BETWEEN COMPARED GROUPS IN MEAN NUMBER OF ADJECTIVES-INFINITIVES PER 10-MINUTE PERIOD, IN THE SETTINGS SPECIFIED

Setting	Out.	In.	Din.	Pict.	All	Setting	Out.	In.	Din.	Pict.	All
	Reg. boys—Rel. boys						Reg. girls—Rel. girls				
Out.	2.37					Out.	7.90				
In.		1.34				In.		2.31			
Din.			0.19			Din.			1.51		
Pict.				2.27		Pict.				3.24	
All					1.61	All					4.42
	Reg. boys—Reg. girls						Rel. boys—Rel. girls				
Out.	7.44					Out.	2.18				
In.		2.54				In.		2.00			
Din.			1.69			Din.			0.00		
Pict.				2.71		Pict.				1.72	
All					4.12	All					1.68
	Reg. boys—Rel. girls						Rel. boys—Reg. girls				
Out.	0.29					Out.	9.22				
In.		0.53				In.		3.88			
Din.			0.20			Din.			1.45		
Pict.				0.55		Pict.				4.88	
All					0.05	All					5.58

pared no differences are statistically reliable, although those for the Outdoor, Picture, Indoor, and combined settings result in critical ratios which indicate approximately 91 or more chances in 100 that true differences exist. Comparisons of these groups as to amount of verbal behavior result in somewhat similar findings (Table 17).

In accordance with the fact that Regular girls have the greatest amount of speech, this group significantly excels all other groups in use of adjectives-infinitives. When Regular girls are compared with Regular boys the critical ratios in the Outdoor and combined settings are indicative of reliable differences, and those in the other settings give approximately 96 or more chances in 100 of the possibility of true differences. Regular girls significantly surpass Relief girls in the Outdoor, Picture, and combined settings, and the critical ratios from the Dinner and Indoor comparisons are indicative, respectively, of 93 and 98.9 chances in 100 of the probability of a true difference. In accordance with the fact that reliable differences exist when the amount of verbal behavior is compared, four of the five comparisons of adjectives-infinitives, as used by Regular girls and Relief boys, result in highly significant differences and the fifth comparison results in a marked difference.

Regular boys and Relief girls have markedly similar records of adjectives-infinitives. It will be recalled that differences are small when these groups are compared on the basis of amount of verbal behavior (Table 17).

Relief boys and Relief girls have *identical records in the Dinner* setting, but other comparisons yield critical ratios which indicate from 96 to 98.6 chances in 100 of the probability of true differences. It should be restated that comparisons of these groups on the basis of the amount of speech result in a statistically reliable difference in the Picture situation and virtually significant differences in all other comparisons.

Comparisons between relatively distinct socio-economic groups with the sex factor constant (Regular boys-Relief boys, and Regular girls-Relief girls) result, respectively, in a marked difference and in a highly reliable difference.

Comparisons between the sexes with the socio-economic factor constant (Regular girls-Regular boys, and Relief girls-Relief boys) result, respectively, in a statistically significant difference, and in a critical ratio which indicates approximately 96 chances in 100 of the possibility of a true difference.

Group differences in incidence of adjectives-infinitives concomitant with differences in sex or in socio-economic status are fairly consistent with the findings obtained from such comparisons of amount of comprehensible speech (Table 17).

Regular girls significantly surpass all other groups in use of adjectives-infinitives. Regular boys and Relief girls have almost identical records, and both markedly excel Relief boys on the basis of the combined settings. The relative standing of the groups in regard to the incidence of adjectives-infinitives is somewhat similar to that for the amount of comprehensible speech. However, the self-comparisons in certain specific settings do not show that use of adjectives-infinitives is affected by change in setting identically as comprehensible vocal output seems to be influenced. It should be noted that more significant differences are concurrent with change in setting for Regular subjects than for Relief subjects. This is consistent with the fact that such critical ratios resulting from comparisons of amount of speech (Table 16) tend to be much larger for Regular subjects than for Relief subjects.

More statistically significant differences in use of adjectives-infinitives result from self-comparisons of subjects in different types of setting than from comparisons of different groups in the same settings. The table of self-comparisons, Table 33, indicates 10 reliable differences, from a possible 24, when correlated observations are used. The table of group comparisons, Table 34, indicates six reliable differences, from a possible 24, when different groups are compared in the same settings. However, if the observations for self-comparisons are treated as if uncorrelated, only three statistically reliable differences result from the self-comparisons in the several settings. On this basis, more significant differences are concurrent with change in socio-economic status or in sex, than with variation in setting.

C. SUMMARY

1. A study of the proportions of the several parts of speech in the total amount of comprehensible verbal behavior for 444 hours, results in the following findings:

(*a*). The two parts of speech which have the largest proportions in the present study are pronouns and verbs, and they maintain approximately the same ratio to the total number of words at all ages. When combined with infinitives they constitute 55 per cent of the speech of the subjects. At all ages they surpass the percentage of nouns.

(*b*). Nouns rank third, with a proportion of 16 per cent, and there is a reduction from 17.9 per cent at 30 months to 13.4 per cent at 54 months. Relief boys have the largest percentage of nouns when compared with the other three groups. Interjections likewise decrease with age, and the mean for the entire group is 2.8 per cent.

(*c*). Articles, conjunctions, prepositions, adverbs, and infinitives increase in varying degrees with age. The proportions are larger for Regular than for Relief subjects, and for girls than boys, except in the case of adverbs.

(*d*). Adjectives do not show a consistent increase with age after a small rise at 36 months. They constitute 7.1 per cent of the verbal behavior of the subjects. Girls have a slightly larger proportion of adjectives than boys, but the difference between the socio-economic groups is negligible.

(*e*). A comparative study of six previous investigations reveals certain trends, such as a decrease of nouns and interjections and an increase in pronouns, adjectives, verbs, and relational words, with age. These trends usually are more marked in those studies which have subjects of 18 and 24 months of age. The present study has subjects above 29 months of age, and the results do not in all cases coincide with those just mentioned.

The six investigations which are cited show marked differences in the proportions of the several parts of speech. The present study shows resemblances in many respects, but likewise differs greatly from certain of the previous studies. It seems probable that many of the differences may be due to the fact that the several investigators have followed different authorities on grammar and the results, strictly speaking, are not directly comparable.

2. The study of the incidence of the parts of speech, or groups of parts of speech, is based upon 345 hours and 20 minutes of observation. The findings are as follows:

(*a*). The reliability coefficients for the parts of speech, or groups of parts of speech, range from .81 to .90 at the various ages, when uniform Outdoor and Indoor records are used.

(*b*). The 120 correlations between the compared settings range from .72 to .93 and indicate, on the part of the specific subjects, a marked degree of consistency from setting to setting in the incidence of the parts of speech, or groups of parts of speech.

(*c*). In accordance with the fact that changes in verbal behavior are concurrent with changes in setting and 22 of 24 self-comparisons result in reliable differences, significant changes in the incidence of the parts of speech, or groups of parts of speech, likewise are concomitant with variations in setting, and 24 self-comparisons in each instance give the following completely reliable differences: nouns, 18; pronouns, 16; verbs, 19; particles, 21; and adjectives-infinitives, 10.

(*d*). The greatest frequencies of the parts of speech, or groups of parts of speech, tend to occur in those settings in which the greatest amounts of verbal behavior are found, that is, in the Outdoor or Picture setting. In the case of Regular girls the Outdoor situation is the scene of the largest records, and in the case of the other three groups the Picture setting usually has the highest incidence.

(*e*). Enumeration, which involves the use of nouns, is especially pronounced in the Picture setting for all groups, although Regular girls frequently have records characterized by some description. The nature of the Picture setting may have been such as to elicit nouns.

(*f*). When groups which differ in socio-economic status or in sex, or in both, are compared on the basis of the combined settings three of six comparisons of verbal behavior result in statistically significant differences. *In accordance with this fact reliable differences* occur when such group comparisons are based upon the parts of speech, or groups of parts of speech, as follows: nouns, two; pronouns, five; verbs, four; particles, three; and adjectives-infinitives, three. Regular girls usually have the largest and Relief boys the poorest records of the four groups. In most cases Relief girls and Regular boys have markedly similar records.

(*g*). When groups which differ in socio-economic status or in sex, or in both, are compared on the basis of verbal behavior in specific rather than combined settings 11 of 24 comparisons result in statistically significant differences. In accordance with this fact reliable differences occur when corresponding group comparisons are based upon the parts of speech, or groups of parts of speech, as follows: nouns, *7;* pronouns, *18;* verbs, *13;* particles, *11;* and adjectives-infinitives, 6.

VI. GENERAL SUMMARY AND CONCLUSIONS

A. Summary

The present study has investigated the variations in language which are concomitant with differences in environmental setting, socio-economic status, sex, age, and other factors. The findings, which are based mainly upon the classification given by Jespersen's *Essentials of English Grammar* (45), have been compared with the results of six other studies of the parts of speech.

Most of the study has been based upon language records which totalled 280 minutes per subject, or 345 hours and 20 minutes for the group of 74 subjects. However, certain parts of the investigation have included 444 hours of recorded verbal behavior.

The subjects range in age from 30 to 65 months, with a mean of 47.6 months. They have been divided into four groups: Regular boys, Relief boys, Regular girls, and Relief girls. Regular subjects came from homes of superior socio-economic levels, and Relief cases from less fortunate families where government aid was being received in some form. Regular and Relief cases had respective percentile ranks of 70 and 46 on the *Merrill-Palmer Scale of Mental* Tests (82).

1. *Length of Sentence*

For Regular boys, Relief boys, Regular girls, and Relief girls in the combined settings the respective means are as follows: 4.56, 3.98, 4.87, and 4.75 words per response.

The length of response shows only slight variations concomitant with change in setting. There are no reliable differences between settings for the combined groups of "All Subjects," the largest critical ratio being 0.86. Only five of the 24 self-comparisons of specific groups in the several settings result in significant differences, and these occur for Relief subjects.

With differences in socio-economic status or in sex noticeable variations occur in length of response. Girls surpass boys and the critical ratio of 2.62 indicates approximately 99.5 chances in 100 of the existence of a true difference. The marked superiority of girls occurs, however, mainly in connection with Relief boys for the comparisons of girls with Regular boys result in unreliable differences. Regular subjects are superior to Relief cases in mean length of response and the critical ratio of 1.76 indicates about 96 chances in 100 of a true difference. The superiority is found particularly in comparisons with Relief boys. All comparisons of Regular subjects with Relief girls result in statistically unreliable differences.

Certain variations in length of response are concomitant with differences in intelligence. Regular subjects, with a percentile rank of 70, surpass Relief cases whose percentile rank is 46. However, marked inconsistencies from comparisons of specific groups indicate that intelligence, as measured in the present study, may not be the basic explanation of the existing differences. A similar statement may be made concerning the fact that Regular subjects, composed of a much larger number of only children and having few siblings as compared with Relief subjects, markedly surpass the latter in mean sentence length.

As age increases a marked increase occurs in the mean length of response. At 30 months the mean response is 3.2 words and at 60 months it has increased to 5.2 words. The proportion of one-word utterances shows a rapid decrease with age.

2. *Number of Words Spoken*

For Regular boys, Relief boys, Regular girls, and Relief girls in the combined settings the comprehensible words spoken per 10-minute period occur with the following frequencies: 46.91, 40.02, 60.77, and 47.27.

The amount of comprehensible speech shows marked variations concomitant with changes in setting. Of 24 self-comparisons of specific groups in the several settings 22 result in highly significant differences. Of the six self-comparisons of "All Subjects" in the different settings five result in completely reliable differences. The Outdoor and Picture settings were the scenes of the greatest amounts of speech.

When compared as to amounts of verbal behavior girls are superior to boys, and Regular subjects to Relief subjects, the differences being statistically reliable. Regular girls surpass each of the other groups and the differences are statistically significant. Relief boys have the smallest records and comparisons with other groups result in differences which are reliable or which closely approach significance.

With differences in intelligence, as indicated by percentile ranks on the *Merrill Palmer Scale of Mental Tests,* marked variations occur in amount of verbal behavior. Regular subjects, with a percentile rank of 70, are significantly superior to Relief subjects whose percentile rank is 46. However, there are several contradictions noted from comparisons of specific groups which indicate that intelligence, as measured in the present study, does not offer a consistent explanation of the existing differences.

Regular subjects, composed of a much larger number of only children and having few siblings as compared with Relief cases, significantly surpass the latter in amount of verbal behavior. Nevertheless, numerous inconsistencies may be noted in connection with the several group comparisons and it is apparent that sibling incidence, if it may be regarded as a possible causative factor, does not exert a uniform influence.

As age increases the amount of speech increases. At 30 months the mean number of comprehensible words per 10-minute period is 20.6, and at 54 months the record is 73.4 words.

Individual records indicate that certain variations in amount of language may be associated with serious personality maladjustments.

3. *Parts of Speech*

A comparative study of six previous investigations reveals certain trends, such as a decrease of nouns and interjections and an increase in pronouns, adjectives, verbs, and certain relational words, with age. Marked differences occur in the proportions of the parts of speech when the results of some of the studies are compared. It appears that a number of the variations may be due to the fact that the investigators followed the viewpoints of different authorities on grammar and, therefore, the results are not directly comparable.

In the present study of the proportions of the parts of speech during 444 hours, pronouns and verbs were found to have the largest percentages. When combined they constitute 55 per cent of the speech of the subjects. At all ages these parts of speech surpass nouns. The respective proportions of nouns, pronouns, verbs, adverbs and adjectives are: 16 per cent, 25.8 per cent, 29.2 per cent, 10.9 per cent, and 7.1 per cent. The respective proportions of conjunctions, prepositions, articles, and interjections are: 1.7 per cent, 4.3 per cent, 2.4 per cent, and 2.8 per cent. Nouns and interjections tend to decrease with age. Articles, conjunctions, prepositions, adverbs, and infinitives increase in varying degrees with age.

In the present study of the incidence of the parts of speech during

345 hours and 20 minutes records were obtained, as previously stated, in four relatively distinct types of environment. The 120 Pearson product-moment correlations between the compared settings range from .72 to .93, indicating a marked degree of consistency on the part of the individual subjects in the incidence of the parts of speech.

Significant changes in the frequencies of the parts of speech are concurrent with changes in setting. From 24 self-comparisons of groups in different settings the following numbers of statistically reliable differences result: nouns, 18; pronouns, 16; verbs, 19; particles, 21; and adjectives-infinitives, 10. These results are more understandable when it is recalled that from 24 similar self-comparisons of the total comprehensible speech used by groups in different settings, 22 statistically significant differences were obtained. The greatest frequencies of the parts of speech, or groups of parts of speech, tend to occur in those settings in which the greatest amounts of verbal behavior are recorded, that is, in the Outdoor or Picture setting.

The records of the several parts of speech vary from group to group. Regular girls usually have the largest and Relief boys the smallest frequencies, and in most instances the records of Relief girls and Regular boys are very similar. When comparisons of these groups are made on the basis of the combined settings statistically significant differences result from the six comparisons as follows: nouns, two; pronouns, five; verbs, four; particles, three; and adjectives-infinitives, three. It should be recalled that three of six such comparisons of verbal behavior result in reliable differences. When the groups are compared on the basis of specific settings rather than combined settings, 11 of 24 comparisons of the amount of speech result in significant differences. In accordance with this fact, reliable differences occur when corresponding group comparisons are based upon the parts of speech, or groups of parts of speech, as follows: nouns, seven; pronouns, 18; verbs, 13; particles, 11; and adjectives-infinitives, six.

B. CONCLUSIONS

The findings of the present investigation seem to warrant the following conclusions:

1. The length of response shows relatively slight variations with change in setting, as the term is used in this study.

2. The length of response varies from group to group. Girls markedly surpass boys and Regular subjects are definitely superior to Relief subjects. Regular girls have the best records and Relief boys the poorest.

3. The length of response shows a consistent increase with age.

4. The amount of comprehensible verbal behavior varies significantly from setting to setting. The Outdoor and Picture situations were the scenes of the largest amounts of speech. The results indicate that in order to secure truly representative samples of the language of preschool subjects it is important to procure records in *several relatively distinct types of environmental situations.* If this is impossible the setting which is used should be carefully described and the limitations recognized.

5. The amount of verbal behavior varies from group to group. Girls are significantly superior to boys, and Regular subjects to Relief subjects. Regular girls surpass all other subjects and Relief boys have the smallest records of any group.

6. The parts of speech show marked variations in frequency concomitant with changes in setting. The changes from setting to setting are, in general, fairly consistent with those occurring for the amount of verbal behavior.

7. The parts of speech vary in incidence from group to group. Regular girls usually have the largest and Relief boys the smallest records.

8. A comparative study of six previous investigations indicates that, although certain common trends are observable, marked differences exist in the proportions of the parts of speech. It appears that an important variable in studies of language development is the use, on the part of investigators, of markedly different classifications in making grammatical analyses. This fact should be recognized as one of the possible causes of the differences in findings whenever such studies are compared.

APPENDIX *A*

DISTRIBUTION OF SUBJECTS

TABLE 35

DISTRIBUTION OF CASES BY AGE, GROUP, SEX, AND PATERNAL OCCUPATION

			Occupational groups*						
Age	Group	Sex	I	II	III	IV	V	VI	VII
30	Reg.	M.	3	1					
	Rel.	M.					1	3	
	Reg.	F.		1	2				
	Rel.	F.					1	1	1
36	Reg.	M.		1	2				
	Rel.	M.						3	
	Reg.	F.	2						
	Rel.	F.						2	
42	Reg.	M.	2	1					
	Rel.	M.					1	2	
	Reg.	F.	1		2				
	Rel.	F.					1	2	
48	Reg.	M.			3				
	Rel.	M.					1	1	1
	Reg.	F.			3				
	Rel.	F.						3	
54	Reg.	M.		1	4				
	Rel.	M.						3	2
	Reg.	F.	1		3				
	Rel.	F.					2	2	
60	Reg.	M.			2				
	Rel.	M.						2	
	Reg.	F.			2				
	Rel.	F.						1	1
All	Reg.	M.	5	4	11	0	0	0	0
	Rel.	M.	0	0	0	0	3	14	3
	Reg.	F.	4	1	12	0	0	0	0
	Rel.	F.	0	0	0	0	4	11	2

*The Occupational Groups are as follows: I. Professional; II. Semi-professional and managerial; III. Clerical, skilled trades, and retail business; IV. Farmers; V. Semi-skilled, minor clerical, and minor business; VI. Slightly skilled trades and occupations requiring little training or ability; VII. Day laborers of all classes (34, pp. 237 and 501).

TABLE 36

MEAN PERCENTILE RANK OBTAINED ON THE MERRILL-PALMER SCALE OF MENTAL TESTS BY AGE, GROUP,* SEX, AND PATERNAL OCCUPATIONS**

Age	Group	Sex	Occupational group** I	II	III	IV	V	VI	VII
30	Reg.	M.	86	70					
	Rel.	M.					75	48	
	Reg.	F.		90	65				
	Rel.	F.					50	40	25
36	Reg.	M.		80	75				
	Rel.	M.						35	
	Reg.	F.	85						
	Rel.	F.						45	
42	Reg.	M.	78	55					
	Rel.	M.					50	40	
	Reg.	F.	70		65				
	Rel.	F.					70	35	
48	Reg.	M.			58				
	Rel.	M.					76	—	35
	Reg.	F.			64				
	Rel.	F.						52	
54	Reg.	M.		80	68				
	Rel.	M.						48	43
	Reg.	F.	80		72				
	Rel.	F.					55	50	
60	Reg.	M.			55				
	Rel.	M.						38	
	Reg.	F.			60				
	Rel.	F.						20	50
	All		80	75	65		63	41	38

*Only 18 Relief boys are included since two did not have mental tests.
**(34, pp. 237 and 501).

Basal Correlations

TABLE 37

Correlations* Between Mean Sentence Lengths Used in the Compared Settings by Regular Boys and Relief Boys

Group	Setting	1	2	3	4	5	6	7
Reg. B.	1 Out.							
	2 In.	.90						
	3 Din.	.92	.91					
	4 Pict.	.95	.93	.95				
Rel. B.	5 Out.				—			
	6 In.					.90		
	7 Din.					.88	.91	
	8 Pict.					.89	.90	.92

*For Regular boys and Relief boys, in this and the succeeding tables, the probable errors of the coefficients between .76 and .81 are .03; of those between .82 and .88 are .02; and of those between .89 and .97 are .01.

TABLE 38

Correlations* Between Mean Sentence Lengths Used in the Compared Settings by Regular Girls and Relief Girls

Group	Setting	1	2	3	4	5	6	7
Reg. G.	1 Out.							
	2 In.	.92						
	3 Din.	.94	.93					
	4 Pict.	.92	.90	.91				
Rel. G.	5 Out.				—	.90		
	6 In.					.93	.94	
	7 Din.					.93	.92	.91
	8 Pict.							

*For Regular girls and Relief girls, in this and the succeeding tables, the probable errors of the coefficients between .77 and .83 are .03; of those between .84 and .90 are .02; and of those between .91 and .96 are .01.

TABLE 39

Correlations Between the Total Number of Comprehensible Words Used in the Compared Settings by Regular Boys and Relief Boys

Group	Setting	1	2	3	4	5	6	7
Reg. B.	1 Out.							
	2 In.	.97						
	3 Din.	.93	.96					
	4 Pict.	.93	.90	.92				
Rel. B.	5 Out.				—			
	6 In.					.93		
	7 Din.					.92	.94	
	8 Pict.					.96	.91	.92

TABLE 40

Correlations Between the Total Number of Comprehensible Words Used in the Compared Settings by Regular Girls and Relief Girls

Group	Setting	1	2	3	4	5	6	7
Reg. G.	1 Out.							
	2 In.	.94						
	3 Din.	.93	.90					
	4 Pict.	.92	.96	.92				
Rel. G.	5 Out.				—			
	6 In.					.95		
	7 Din.					.93	.93	
	8 Pict.					.91	.91	.92

TABLE 41

Correlations Between Nouns Used in the Compared Settings by Regular Boys and Relief Boys

Group	Setting	1	2	3	4	5	6	7
Reg. B.	1 Out.							
	2 In.	.86						
	3 Din.	.82	.86					
	4 Pict.	.85	.89	.85				
Rel. B.	5 Out.				—			
	6 In.					.89		
	7 Din.					.81	.79	
	8 Pict.					.81	.76	.73

TABLE 42

Correlations Between Nouns Used in the Compared Settings by Regular Girls and Relief Girls

Group	Setting	1	2	3	4	5	6	7
Reg. G.	1 Out.							
	2 In.	.89						
	3 Din.	.82	.79					
	4 Pict.	.75	.80	.78				
Rel. G.	5 Out.				—			
	6 In.					.81		
	7 Din.					.81	.78	
	8 Pict.					.80	.83	.82

TABLE 43

Correlations Between Pronouns Used in the Compared Settings by Regular Boys and Relief Boys

Group	Setting	1	2	3	4	5	6	7
Reg. B.	1 Out.							
	2 In.	.86						
	3 Din.	.82	.87					
	4 Pict.	.86	.87	.83				
Rel. B.	5 Out.				—			
	6 In.					.83		
	7 Din.					.77	.77	
	8 Pict.					.85	.80	.72

TABLE 44

Correlations Between Pronouns Used in the Compared Settings by Regular Girls and Relief Girls

Group	Setting	1	2	3	4	5	6	7
Reg. G.	1 Out.							
	2 In.	.89						
	3 Din.	.86	.87					
	4 Pict.	.87	.90	.84				
Rel. G.	5 Out.				—			
	6 In.					.82		
	7 Din.					.81	.80	
	8 Pict.					.82	.86	.86

TABLE 45

Correlations Between Verbs Used in the Compared Settings by Regular Boys and Relief Boys

Group	Setting	1	2	3	4	5	6	7
Reg. B.	1 Out.							
	2 In.	.93						
	3 Din.	.92	.86					
	4 Pict.	.87	.88	.32				
Rel. B.	5 Out.				—			
	6 In.					.85		
	7 Din.					.80	.83	
	8 Pict.					.35	.84	.93

TABLE 46

Correlations Between Verbs Used in the Compared Settings by Regular Girls and Relief Girls

Group	Setting	1	2	3	4	5	6	7
Reg. G.	1 Out.							
	2 In.	.86						
	3 Din.	.77	.77					
	4 Pict.	.87	.78	.73				
Rel. G.	5 Out.				—			
	6 In.					.90		
	7 Din.					.84	.85	
	8 Pict.					.88	.87	.78

TABLE 47

Correlations Between Particles Used in the Compared Settings by Regular Boys and Relief Boys

Group	Setting	1	2	3	4	5	6	7
Reg. B.	1 Out.							
	2 In.	.81						
	3 Din.	.75	.82					
	4 Pict.	.86	.89	.83				
Rel. B.	5 Out.				—			
	6 In.					.76		
	7 Din.					.78	.84	
	8 Pict.					.81	.85	.86

TABLE 48

Correlations Between Particles Used in the Compared Settings by Regular Girls and Relief Girls

Group	Setting	1	2	3	4	5	6	7
Reg. G.	1 Out.							
	2 In.	.78						
	3 Din.	.80	.75					
	4 Pict.	.89	.79	.81				
Rel. G.	5 Out.				—			
	6 In.					.79		
	7 Din.					.78	.86	
	8 Pict.					.80	.88	.86

TABLE 49

CORRELATIONS BETWEEN ADJECTIVES-INFINITIVES USED IN THE COMPARED SETTINGS BY REGULAR BOYS AND RELIEF BOYS

Group	Setting	1	2	3	4	5	6	7
Reg. B.	1 Out.							
	2 In.	.82						
	3 Din.	.85	.90					
	4 Pict.	.81	.83	.76				
Rel. B.	5 Out.				—			
	6 In.					.85		
	7 Din.					.78	.76	
	8 Pict.					.77	.80	.79

TABLE 50

CORRELATIONS BETWEEN ADJECTIVES-INFINITIVES USED IN THE COMPARED SETTINGS BY REGULAR GIRLS AND RELIEF GIRLS

Group	Setting	1	2	3	4	5	6	7
Reg. G.	1 Out.							
	2 In.	.85						
	3 Din.	.84	.86					
	4 Pict.	.85	.83	.83				
Rel. G.	5 Out.				—			
	6 In.					.83		
	7 Din.					.77	.77	
	8 Pict.					.80	.78	.81

REFERENCES

Anderson, J. E. (Chr.) The Young Child in the Home: A survey of three thousand familes. (Report of the committee on the infant and preschool child, White House conference on child health and protection.) New York: Appleton-Century, *1936.* *Pp.* xxi + *415.*

————. An evaluation of various indices of linguistic development. *Child Devel.,* 1937, **8**, 62-68.

Arrington, R. E. Interrelations in the behavior of young children. *Teach. Coll. Child Devel. Monog.,* 1932, No. 8. Pp. xvii + 156.

Bateman, W. G. Two children's progress in speech. *J. Educ. Psychol.,* 1915, **6**, 475-493.

Bayley, N. A cumulative study of environmental correlates of intelligence. *Psychol. Bull., 1935,* **32**, *702.*

Blackfan, K. D. (Chr.) Growth and Development of the Child. Part IV. Appraisement of the child. (Report of the committee on growth and development, White House conference on child health and protection.) New York: Century, 1932. Pp. xix + 344.

Boyd, W. The development of a child's vocabulary. *Ped. Sem.,* 1914, **21**, 95-124.

————. The development of sentence structure in childhood. *Brit. J. Psychol., 1926-27,* **17**, *181-191.*

Boynton, P. L. Intelligence, Its Manifestations and Measurement. New York: Appleton, 1933. Pp. xi + 446.

Brandenburg, G. C., & Brandenburg, J. Language development during the fourth year: The conversation. *Ped. Sem.,* 1919, **26**, 27-40.

Bridges, K. M. B. Critical notes on mental tests for children of preschool age. *Ped. Sem.,* 1927, **34**, 38-44.

Brooks, F. D., & Shaffer, L. F. Child Psychology. Boston: Houghton Mifflin, *1937.* Pp. xxx + *600.*

Coffey, H. S., & Wellman, B. L. The rôle of cultural status in intelligence changes of preschool children. *J. Exper. Educ.,* 1936-37, **5**, 191-202.

Cooke, D. H. Minimum Essentials of Statistics. New York: Macmillan, 1936. Pp. xviii + 271.

Curme, G. O. Parts of Speech and Accidence. Boston: Heath, 1935. Pp. xiii + 370.

Davis, E. A. The development of linguistic skill in twins, singletons with siblings, and only children from age five to ten years. *Univ. Minn. Inst. Child Welf. Monog. Ser.,* 1937, No. 14. Pp. x + 165.

————. Mean sentence length compared with long and short sentences as a reliable measure of language development. *Child Devel.,* 1937, **8**, 69-79.

————. The mental and linguistic superiority of only girls. *Child Devel.,* 1937, **8**, 138-143.

Day, E. J. The development of language in twins. *Child Devel.,* 1932, **3**, 179-199, 298-316.

Descoeudres, A. Le Développment de L'Enfant de Deux à Sept Ans. Neuchâtel and Paris: Delachaux & Niestlé, *1921.* *Pp. 325.*

21. Dewey, J. The psychology of infant language. *Psychol. Rev.*, 1894, **1**, 63-66.

22. Drever, J. A study of children's vocabularies: I and II. *J. Exper. Ped.*, 1915-16, **3**, 34-43, 96-103, 182-188.

23. Driscoll, G. P. The developmental status of the preschool child as a prognosis of future development. *Teach. Coll. Child Devel. Monog.*, 1933, No. 13. Pp. xiv + 111.

24. Edwards, A. S. A theoretical and clinical study of so-called normality. *J. Abn. & Soc. Psychol.*, 1933-34, **28**, 366-376.

25. Ellesor, M. V. The relation between situation and response in vocalization of a three year old child. *Child. Devel.*, 1934, **5**, 158-164.

26. Fales, E. A comparison of the vigorousness of play activities of preschool boys and girls. *Child Devel.*, 1937, **8**, 144-158.

27. Fisher, M. S. Language patterns of preschool children. *Teach Coll. Child. Devel. Monog.*, 1934, No. 15. Pp. xvi + 88.

28. Garrett, H. E. Statistics in Psychology and Education. New York: Longmans, Green, 1937. Pp. xiv + 493.

29. Garrison, S. C., & Garrison, K. C. Fundamentals of Psychology in Secondary Education. New York: Prentice-Hall, 1936. Pp. xx + 599.

30. Gesell, A., & Lord, E. E. A psychological comparison of nursery school children from homes of low and high economic status. *Ped. Sem.*, 1927, **34**, 339-356.

31. Goodenough, F. L. Measuring behavior traits by means of repeated short samples. *J. Juv. Res.*, 1938, **12**, 230-235.

32. ———. Inter-relationships in the behavior of young children. *Child Devel.*, 1930, **1**, 29-47.

33. ———. The measurement of mental growth. In *A Handbook of Child Psychology*, 2nd ed., ed. by C. Murchison. Worcester, Mass.: Clark Univ. Press, 1933. (Pp. 303-328.)

34. Goodenough, F. L., & Anderson, J. E. Experimental Child Study. New York: Century, 1931. Pp. xii + 546.

35. Goodenough, F. L., & Leahy, A. M. The effect of certain family relationships upon the development of personality. *J. Genet. Psychol.*, 1927, **34**, 45-71.

36. Grant, J. R. A child's vocabulary and its growth. *Ped. Sem.*, 1915, **22**, 183-203.

37. Hattwick, B. W. The influence of nursery school attendance upon the behavior and personality of the preschool child. *J. Exper. Educ.*, 1936-37, **5**, 180-190.

38. Hildreth, G. Occupational status and intelligence. *Person. J.*, 1934-35, **13**, 153-157.

39. Horn, E. The commonest words in the spoken vocabulary of children up to and including six years of age. *Yearb. Nat. Soc. Stud. Educ.*, 1925, **24**, Part 1, 185-199.

40. Jespersen, O. Language, its Nature, Development and Origin. London: Allen & Unwin, 1922. Pp. 448.

41. ———. The Philosophy of Grammar. London: Allen & Unwin, 1924. Pp. 359.

42. ———. A Modern English Grammar. Heidelberg: Carl Winters Universitatsbuchhandlung, 1927, 1, Part II. Pp. xxviii + 512.

43. ———. A Modern English Grammar. Heidelberg: Carl Winters Universitatsbuchhandlung, 1927, 2, Part III. Pp. ix + 415.

44. ———. A Modern English Grammar. Heidelberg: Carl Winters Universitatsbuchhandlung, 1931, 3, Part IV. Pp. xxxi + 400.

✓45. ———. Essentials of English Grammar. New York: Holt, 1933. Pp. 387.

46. Joint Committee on Grammatical Nomenclature. Report. *Nat. Educ. Assoc.*, 1923. Pp. x + 75.

47. Jones, H. E. Relationships in physical and mental development. *Rev. Educ. Res.*, 1936, **6**, 102-123.

48. Jones, M. C., & Burks, B. S. Personality development in childhood—a survey of problems, methods and experimental findings. *Monog. Soc. Res. Child Devel.*, 1936, **1**, No. 4. Pp. vi + 205.

49. Jordon, A. M. Parental occupations and children's intelligence scores. *J. Appl. Psychol.*, 1933, **17**, 103-119.

50. Kantor, J. R. An Objective Psychology of Grammar. Bloomington: Ind. Univ., 1936. Pp. xvi + 344.

51. Kawin, E., & Hoefer, C. A Comparative Study of a Nursery-School Versus a Non-Nursery-School Group. Chicago: Univ. Chicago Press, 1931. Pp. ix + 52.

52. Kirkpatrick, E. A. Fundamentals of Child Study. New York: Macmillan, 1917. Pp. xxiii + 380.

53. Lewis, M. M. Infant Speech: A Study of the Beginnings of Language. New York: Harcourt, Brace, 1936. Pp. xii + 335.

54. Little, M. F., & Williams, H. M. An analytical scale of language achievement. *Univ. Iowa Stud.: Stud. Child Welf.*, 1937, **13**, No. 2, Part 4. Pp. 49-77.

55. Low, A. A. Studies in infant speech and thought: Part I. The development of sentence structure in infancy from the standpoint of grammar. *Univ. Ill. Bull.*, 1936, **33**, No. 39. Pp. 71.

56. McCarthy, D. A comparison of children's language in different situations and its relation to personality traits. *J. Genet. Psychol.*, 1929, **36**, 583-591.

✓57. ———. The language development of the preschool child. *Univ. Minn. Inst. Child Welf. Monog. Ser.*, 1930, No. 4. Pp. xiii + 174.

58. ———. Language development. In *A Handbook of Child Psychology*, 1st ed., ed. by C. Murchison. Worcester, Mass.: Clark Univ. Press, 1931. (Pp. 278-315.)

59. ———. Language development. In *A Handbook of Child Psychology*, 2nd ed., ed. by C. Murchison. Worcester, Mass.: Clark Univ. Press, 1933. (Pp. 329-373.)

60. Maller, J. B. Mental ability and its relation to physical health and social economic status. *Psychol. Clin.*, 1933-34, **22**, 101-107.

61. Malley, H. Growth in social behavior and mental activity after six months in nursery school. *Child Devel.*, 1935, **6**, 303-309.

62. Markey, F. V. Imaginative behavior of preschool children. *Teach. Coll. Child Devel. Monog.*, 1935, No. 18. Pp. xvi + 139.

63. Markey, J. F. The Symbolic Process and its Integration in Children. New York: Harcourt, Brace, 1928. Pp. xii + 192.

64. Mateer, F. The vocabulary of a four-year-old boy. *Ped. Sem.*, 1908, **15**, 63-74.

65. Mowrer, W. M. C. Performance of children in Stutsman tests. *Child Devel.*, 1934, **5**, 93-96.

66. National Society for the Study of Education. Nature and nurture: Their influence upon achievement. *Yearb. Nat. Soc. Stud. Educ.*, 1938, **27**, Part II, 268-282.

✓67. Nice, M. M. Concerning all day conversations. *Ped. Sem.*, 1920, **27**, 166-177.

68. ———. Length of sentences as a criterion of a child's progress in speech. *J. Educ. Psychol.*, 1925, **16**, 370-379.

69. ———. Analysis of conversation of children and adults. *Child Devel.*, 1932, **3**, 240-246.

70. ———. A child's attainment of the sentence. *J. Genet. Psychol.*, 1933, **42**, 216-224.

✓71. Olson, W. C., & Koetzle, V. S. Amount and rate of talking of young children. *J. Exper. Educ.*, 1936-37, **5**, 175-179.

72. Palmer, C. Height and weight of the depression poor. *U. S. Pub. Health Rep.*, 1935, **50**, 1106-1113.

73. Pintner, R. Intelligence Testing: Methods and Results. (2nd ed.) New York: Holt, 1931. Pp. xii + 555.

74. Portenier, L. G. Factors influencing the social adjustment of children of preschool age. *J. Genet. Psychol.*, 1937, **51**, 127-139.

75. Robinson, E. W., & Conrad, H. S. The reliability of observations of talkativeness and social contact among nursery school children by the short time sample technique. *J. Exper. Educ.*, 1933-34, **2**, 161-165.

76. Rundquist, E. A., & Sletto, R. F. Personality in the Depression: A Study in the Measurement of Attitudes. *Univ. Minn. Inst. Child. Welf. Monog. Ser.*, 1936, No. 12. Pp. xxii + 398.

77. Seth, G., & Guthrie, D. Speech in Childhood: Its Development and Disorders. London: Oxford Univ. Press, 1935. Pp. x + 224.

78. Smith, M. E. An investigation of the development of the sentence and the extent of vocabulary in young children. *Univ. Iowa Stud.: Stud. Child Welf.*, 1926, **3**, No. 5. Pp. 92.

✓79. ———. A study of some factors influencing the development of the sentence in preschool children. *J. Genet. Psychol.*, 1935, **46**, 182-212.

80. Stalnaker, E. Language of the preschool child. *Child Devel.*, 1933, **4**, 229-236.

✓81. Stern, W. Psychology of Early Childhood up to the Sixth Year of Age. (Trans. from the 3rd German ed. rev. and enl.—by A. Barwell.) London: Allen & Unwin, 1924. Pp. 557.

✓82. Stutsman, R. Mental Measurement of Preschool Children. New York: World Book, 1931. Pp. x + 368.

83. Terman, L. M. The Measurement of Intelligence. Boston: Houghton Mifflin, 1916. Pp. xviii + 362.

84. Tracy, F., & Stimpfl, J. Psychology of Childhood. Boston: Heath, 1909. Pp. x + 219.

85. Van Alstyne, D. The environment of three-year-old children: Factors related to intelligence and vocabulary tests. *Teach. Coll. Contrib. Educ.*, 1929, No. 366. Pp. vii + 108.

86. Waddle, C. W. An Introduction to Child Psychology. Boston: Houghton Mifflin, 1918. Pp. xv + 317.

87. Wellman, B. L. Some new bases for interpretation of the *IQ*. *J. Genet. Psychol.*, 1932, **41**, 116-126.

88. ———. The effect of pre-school attendance upon the *IQ*. *J. Exper. Educ.*, 1932-33, **1**, 48-69.

89. ———. Growth in intelligence under differing school environments. *J. Exper. Educ.*, 1934-35, **3**, 59-83.

✓ 90. Williams, H. M. An analytical study of language achievement in preschool children. *Univ. Iowa Stud.*: *Stud. Child Welf.*, 1937, **13**, No 2, Part I. Pp. 18.

91. Williams, H. M., & McFarland, M. L. A revision of the Smith vocabulary test for preschool children. *Univ. Iowa Stud.*: *Stud. Child Welf.*, 1937, **13**, No. 2, Part III, 35-46.

92. Zyve, C. T. Conversation among children. *Teach. Coll. Rec.*, 1927, **29**, 46-61.

Published as a separate and in *Genetic Psychology Monographs*, 1941, **23**, 143-189.

INFANT DEVELOPMENT UNDER CONDITIONS OF RESTRICTED PRACTICE AND OF MINIMUM SOCIAL STIMULATION*

Department of Psychology, University of Virginia

WAYNE DENNIS

Acknowledgments 145

I. Introduction 147

II. Subjects and experimental conditions 149

III. A biographical account of the subjects' behavior 155

IV. A list of developmental items 173

V. Discussion 179

VI. Summary 187

References 189

ACKNOWLEDGMENTS

We wish to express our appreciation to the Institute for Research in the Social Sciences at the University of Virginia for financial support of the research here reported, to Mrs. Dennis for her invaluable collaboration, and to Dr. L. T. Royster and Miss Rose Williams of the University of Virginia Hospital for their assistance and cooperation. The Research Committee of the University of Virginia has defrayed the cost of publication.

WAYNE DENNIS

Department of Psychology
University of Virginia
Charlottesville, Virginia

I. INTRODUCTION

Some years ago we recorded the behavioral development of two infants who were reared from the beginning of the second to the end of the fourteenth month of life under conditions which may be designated as restricted practice and minimum social stimulation. The essentials of these conditions were as follows: The two experimenters (Mrs. Dennis and the writer) exercised the sole care of the subjects. The subjects saw no other children and saw very few adults aside from the experimenters. We refrained from performing any infant responses in the subjects' presence. We did not reward or punish the subjects for any response nor instruct them in any way. Furthermore, the environment of the children was such that in respect to many responses there was little or no opportunity to engage in practice. The aim of the study was to determine what behavioral development would occur when the care of the infants was reduced to the minimum attention which would insure their comfort and well-being.

Such an investigation necessarily gathered data bearing upon many aspects of infant behavior. We have previously reported certain portions of our data. In one of these reports (6) it has been shown that our subjects exhibited all of the responses which biographical studies (7) and Shirley's investigation (10) have shown to be common responses of infants, and that throughout most of the first year our subjects' responses occurred within the usual age range of each response. In the period following the first birthday one subject showed retardation in regard to the onset of locomotor items because of a partial hemiplegia which had not been noticed at an earlier period. This disability was probably referable to a cranial injury at birth, and could in no way be attributed to the experiment. The second subject, in respect to locomotor development, was near, but not entirely beyond, the upper extreme of control cases. A second report (2) revealed that both subjects were retarded in respect to a few responses the practice of which was entirely ruled out by the conditions of the experiment. These responses were: sitting without support, reaching for a dangling object, and supporting the weight of the body upon the feet. While these responses were absent when they were tested for at ages beyond the usual period for their appearance, they were readily established when opportunities for practice were afforded to the subjects.

A further phase of the investigation had to do with smiling to social stimulation (1). It was found that visual stimulation from the experimenters caused smiling, although the experimenters had not fondled or petted the infants. This appeared to be because the visual appearance of adults had been associated with the feeding, bathing, cleaning, and dressing of the infants. This interpretation was corroborated by the fact that early in the experiment the speech of the experimenters, in contrast to the visible presence of the experimenters, did not cause smiling. Smiling upon verbal stimulation occurred soon after the investigators began to announce their entrance to the nursery and to speak to the subjects while attending to their wants.

Handedness, and other forms of laterality of function (3), were exhibited by the subjects despite the fact that we attempted to treat the two sides of the body in identical ways. The handedness of the subjects differed, although the two infants were subjected to identical environmental conditions.

Other portions of our data await further report. The delay in the presentation of the findings has arisen in part from intervening activities which required that the analysis of the data be postponed and in part from the time-consuming nature of the analysis of the data, especially of those data which were obtained by the use of the motion picture camera.

The aim of the present monograph is to present a general picture of the month-by-month development of the subjects. Such an account will provide a survey of the behavior of the subjects which cannot be obtained by a combination of the studies of special responses. No method has yet replaced the biographical method as a scheme for glimpsing the development of the infant as a series of related events.

II. SUBJECTS AND EXPERIMENTAL CONDITIONS

In earlier communications we have described our subjects and have depicted the experimental conditions as they bore upon the special topics upon which we were reporting. It seems desirable here to present the conditions of the experiment in their entirety. We hope that the account is sufficiently complete so that no further account of experimental conditions need be given at a future time.

Our subjects were fraternal twins, but the relationship of the infants was not an important part of the study. Our original plan envisaged the use of only one subject, but the first inquiry for a subject to the Social Service Department of the University of Virginia Hospital revealed an opportunity to secure a pair of non-identical female twins. The proffered twins were accepted because they enabled us to double the number of our subjects with considerably less than a doubling of cost and care. Thus the twin relationship of the subjects was, from the standpoint of our problem, purely accidental.

Their parents are of north-European ancestry. The father was a former taxi-driver, whose whereabouts at the time of the birth of the twins was unknown. The mother had been a saleswoman in department stores. The father had no other children; the mother had two children by a former husband. We were able to obtain the twins as subjects because the mother was unable to provide for them. They came to us when they were 36 days old. The mother understood that we offered temporary care of the twins in return for the privilege of studying them. She understood the general nature of our research and was coöperative at all times.

Throughout the experiment the twins lived in our home but they were confined to the nursery. This was a second-floor room, so situated that from the infants' position only sky and tree tops were visible through the windows. The room itself contained only the subjects' cribs, a bureau, a table, two chairs, and a screen near the door. No picture or decoration of any sort was permitted in the nursery. The door of the room was kept closed, and we entered the room only to care for the subjects, to observe them, and to experiment with them. The twins were never taken from the nursery except on a few test occasions and for trips to the University Hospital for anthropometric and pediatric examination. On these trips the infants' faces were covered. For three months during the summer the subjects lived in a summer cabin, 17 miles away, in a room similar to the winter nursery. As the trip to and from the cabin was carefully supervised, and as their mode of life at the cabin was identical with that in town, this fact does not alter the principle that the environment of the subjects was limited and controlled in a very strict manner.

Although the infants were indoors at all times, we saw to it that fresh air and sunshine were plentifully supplied. Windows were opened for a period daily, the length of the period depending upon the outside temperature. The subjects, dressed only in shirt and diaper, were given sun baths before an open window. Direct sunlight came through the windows during a large part of the daylight hours. In spite of spending the first year of life indoors, the twins had a sun tan at all times.

The subjects were placed in individual cribs, of the trade name "Kiddie Koop." The cribs were placed side by side with a screen equal in height to the cribs between the two, so that the twins could see each other only when taken from their beds. During the first nine months the subjects were taken from the cribs only for feeding and bathing or when removal from the cribs was demanded for the purposes of experimentation.

With the exception of a few occasions during the latter part of the experiment the sole care of the twins was supplied by the experimenters. This means that we bathed and fed the infants, changed the diapers and bed clothing, and cleaned the room. The infants seldom saw other people, and when they did it was with our knowledge and supervision. Visitors were required to adhere to the same practices which we imposed upon ourselves.

We carefully refrained from rewarding or punishing the subjects for any action. The subjects were not scolded or spanked or treated roughly in connection with any response. We did conduct some experiments to determine the first reaction of the subjects to scolding and to a light slap on the thigh, but these stimuli came at a time when the infants were silently and quietly watching the experimenter. With a few exceptions, we never encouraged or discouraged any act of the twins. The exceptions to this rule, and to other such general rules, occurred in the last month of the investigation, when the experimental conditions were partially suspended.

Indifferent behavior toward infants is most difficult to achieve, not in respect to avoiding such overt acts as praising and fondling but in respect to the amount of attention which is given to the infant. If the infant is doing something new or interesting, the adult is likely to watch the child much more closely than usual, and this increased attentiveness on the part of the adult may act as a reward and may induce repetition of the response. In order to avoid variation in attentiveness we attempted to watch the infants at all times when we were in their presence, except when we were taking notes. In this connection it must be borne in mind that when we were in the nursery we were never engaged in anything other than meeting the needs of the subjects and observing their behavior. Routine behavior, as well as any new behavior of the subjects, was observed and recorded each time we entered the nursery. This is very different from the ordinary home situation, where the infant receives attention primarily when he engages in some unusual, or some culturally significant, response. It is our belief that we did not reveal more interest in any one act of our subjects than we did in any other act.

This leads to a consideration of our treatment of the twins' cries. In this respect our procedure was as follows: If either of the infants cried at some time other than feeding time, one of the experimenters entered the room and investigated the cause. The cause was ordinarily a soiled diaper. The diaper was changed, and we left the room. *If the child cried as we left, we paid no attention to such a cry.*

The feeding routine was maintained in a general way, but not with extreme rigidity. If an infant was asleep at the regular time for feeding, we allowed her to go somewhat beyond the usual period. If she fretted a great deal or cried strongly before the hour for feeding, her food was given as much as 15 minutes early. In that sense, we may be said to have rewarded the crying of the subjects, but in so doing we could hardly have rewarded any other response. We shall see later that crying was very infrequent.

We not only avoided reward and punishment but we avoided acts which might have provided examples for imitation. With certain exceptions to be noted later our behavior in the nursery was limited to changing diapers, bathing, feeding, etc. We carefully refrained from baby talk and from babbling, as we wanted to know whether such vocalizations would occur without example. Likewise, we never performed for the twins such acts as patting their hands or playing with their toes.

Thus far we have spoken only of the conditions which remained relatively constant until the last few weeks of the experiment. We turn now to more stringent restrictions in the environment of the subjects which, in the main, were applied only for the first half-year. The conditions to be described were designed to provide answers to specific questions and were abandoned when the answers were obtained.

We wished to determine whether or not the infants would smile upon hearing the voice of the adult, if speech were not associated with the care and attention which the adult supplied. For this reason, until the twins were 26 weeks of age, we never announced our entry into the nursery and never spoke to the subjects. We were not totally silent, for we occasionally commented to each other while in the nursery, but we were careful not to make comments while we were feeding or otherwise caring for the twins. Our speech when we were outside the nursery could be heard by the infants, but it had no more relation to their behavior than did traffic noises or other common sounds.

We wished also to know whether positive responses toward us would develop if we refrained from smiling at the twins and from petting, cuddling, and fondling them. In order to determine the answer to this question we avoided these expressions during the first 26 weeks. Withholding of demonstrations of affection of this sort was not an easy task to impose upon ourselves, particularly as the subjects themselves were very expressive. From the 15th week onward they almost invariably greeted us with a smile and a vocalization. After this fact was thoroughly established, we decided in Week 27 to return their smile of greeting, and to speak to them as we approached.

The reader may ask how we could keep from rewarding various responses of the subjects when once we permitted ourselves to smile at and to fondle and play with the twins. This was accomplished by engaging in these responses only at times when no particular act of the twins could be said to be encouraged by our behavior. We romped with the children, as by shaking or rolling them only when they were indifferently employed with some response that was already well-established.

During the 36th week we began placing the twins for several minutes each day upon quilted pads on the floor. This was for the purpose of giving them some experience on a relatively hard surface. At the beginning of the 42nd week we started placing the infants in high-chairs for a short period each day. But even in the last month of the experiment, the larger part of the twins' day was spent in the cribs.

No toys were provided for the subjects until the 49th week. At that time a rattle and a few other toys, to be described later, were introduced. Our aim in introducing the toys was to see what use would be made of them when no example, encouragement, or instruction was given by the adult.

One sort of restriction to which the infants were subjected requires special comment. In accordance with our general policy of letting the twins alone and of refraining from initiating any development, we did not put the twins into sitting or standing postures, nor provide any practice in reaching for many months. Under normal circumstances parents prop the child up before he can sit alone, hold him upright on the adult's lap before he can stand well, and dangle objects before him before he can reach with any accuracy. These responses, more than any others, require the intervention of the adult in order that early practice be made possible. Our experiment ruled out nearly all possibilities of early practice of these reactions, whereas many other responses were practiced by the twins because unlike the responses mentioned above they did not require the aid of the adult.

The three responses—sitting alone, standing with support, and reaching for a dangling object—were submitted to tests when the twins were of an age such that the responses are present in all normal infants of normal environment. The results have been dealt with in an earlier publication (2). When these tests had been made, we no longer refrained from placing the twins in sitting and standing postures, nor did we keep all movable objects out of their reach.

We attempted to give the two subjects identical treatment. This was carried to the extent of dividing the attention of each experimenter equally between the two subjects. For convenience, the experimenters will be referred to by their initials: *W* referring to the writer and *M* referring to Mrs. Dennis. *M* fed Rey throughout one day while *W* attended to the feeding of Del. During the next day, *M* fed Del and Rey was fed by *W*, etc. The infants were bathed one at a time, but the order of bathing was reversed from day to day. The bathing was done by *W*, the drying and dressing by *M*.

In order that the infants' positions in the room should not lead to a difference in treatment, the crib which Rey occupied during one day was occupied by Del during the next day, etc. Thus it will be seen that great pains were taken to insure that the two subjects were treated exactly alike.

A few words may be added on miscellaneous aspects of the twins' care. Before the twins came to us, the mother's milk had been found to be quite inadequate, and supplementary feeding by bottle had been resorted to. At the beginning of the experiment, bottle feeding constituted the sole means of feeding. The nursing bottle was utilized for the giving of liquid foods throughout the experiment, i.e., weaning from the bottle was not attempted, but solid and semisolid foods were given by spoon.

The diet of the subjects was chosen with the advice of Dr. L. T. Royster of the Pediatrics Department of the University of Virginia Hospital. During the early months the diet of the twins consisted of lactic acid milk, which was kindly prepared for us by the Hospital staff, and of orange juice. Soup, cereal, and strained fruits and vegetables were introduced gradually after the 24th week. No elimination training of any kind was undertaken.

III. BIOGRAPHICAL ACCOUNT OF THE SUBJECTS' BEHAVIOR

The major part of our records consist of day-by-day observations of behavior which fill about 1,800 pages of note books. Notes were taken nearly every time either experimenter entered the room to care for the babies. In addition to entrances into the room which were made in order to care for the infants and to experiment with them, several entrances per day were made solely to note down the *condition and activities* of the babies. *Habitual* as well as *new* responses were recorded. On the average, notes were written 11.88 times per day. A feeding period or an experiment, which may have lasted for half an hour or more and during which notes may have been taken continuously or on several occasions, was counted as only one note-taking in computing this average. Allowing 12 hours of undisturbed night sleep for the babies, we obtained, on the average, one record of activity during each of the remaining hours. We estimate that one or both experimenters were in the room during about two hours per day. As a corollary it should be pointed out that this means that the babies were not "isolated."

The notebooks contain observations not only of the spontaneous activities of the twins but also records of their reactions in test situations. These notes include behavior upon first being placed in a sitting position, and upon first being given an opportunity to stand,—tests which were mentioned above—and also behavior in a number of other tests which will be referred to in later sections.

Some of these tests consisted in stimuli of short duration, such as stimulation by odors, by sounds, by pricking with a pin, and by tickling. Many of these were grouped into series of stimuli, the members of the series always being given in the same order. These are called the Regular Series. As a rule, each series was presented once per day for 20 days. Since many of these experiments did not yield results of any significance, not all of them are included in the present report. However, this material is in the author's files and is available to anyone who wishes to consult it.

The results of two groups of these experiments we hope to present in some detail at a later date. These experiments constituted two series which were recorded by motion picture camera. The analysis of the films is so extensive a task that its presentation will require a special report. Those aspects of the moving picture records which are most relevant to our present problems will be referred to at appropriate points.

We planned and utilized many test situations in the course of the experiment, and these greatly increased the labor of the research. Nevertheless we feel in retrospect that the day-by-day observations of the subjects proved to be of more significance than did the test situations. It will be found that the account which follows is based chiefly upon the routine daily observations. This is not because the results of the test situations were ignored but because in most instances the tests elicited only behavior which had already been witnessed in the course of our day-by-day contact with the twins.

We turn now to a series of monthly summaries of the twins' development.

A. Second Lunar Month (Weeks 6-8)

When the twins came under our care at 36 days of age, they were not healthy. They weighed less than they had weighed at birth and their birth weights were such as nearly to place them in the category of prematures.[1]

Although they were in a malnourished condition, they would take only three or four ounces of milk at a feeding, a fact which was due to irregular feeding habits. Our first concern was to establish a four-hour feeding routine. The four-hour interval between feedings at first caused a certain amount of crying, but within four days the infants were well-adapted to the new schedule, seldom awakened between feeding periods and consumed much more milk at each feeding than formerly. They began to gain in weight immediately. During the first month of the experiment each subject *doubled* her weight, and for the remainder of the experiment each child was of good size and weight for her age.

At the beginning of the experiment, both babies were inflamed and sore in the anogenital region. Careful attention to cleanliness, and the application of olive oil and a soothing powder, corrected this condition within the first two days. The navel of each infant was slightly protuberant when the children entered the nursery. It was necessary to keep the navel under adhesive tape for several weeks before this condition was remedied.

By the end of the second week of the experiment, the twins could be said to have been in general good health. With the exception of two slight colds they remained in good health throughout the course of the investigation.

---

[1] Birth weights: Rey, 2455 grams; Del, 2585 grams.

In the sixth week, when the experiment began, the infants' behavior could not be distinguished from the responses of newborn infants. They did not look at us nor did they focus upon any still or moving objects. The line of regard was mainly in a forward direction with but little shifting. The twins did not smile when fed nor when we entered the room; the occasional smiles which we saw seemed unrelated to anything which was being done to the subjects.

There was a small amount of random movement, but the subjects were relatively inactive. Vocalization was rare. The babies stretched and yawned, ate and slept.

In the seventh week, definite advances were observed. The infants looked about, began to follow us with their eyes, and smiled occasionally upon seeing us. From the time that the infants first showed visual regard for adults, they were most attentive to *faces* rather than to other parts of the persons at whom they were looking.

The infants, of course, cried when hungry from the beginning, but some changes were soon noticed in this crying. In the earliest weeks, hunger crying did not stop until the nipple was placed in the child's mouth. Later the activity of the adult indicated to each child that she was to be fed, and crying ceased before feeding had commenced. It was specifically our entrance into the nursery that began, in the eighth week, to cause the hungry subjects to cease crying.

B. Third Lunar Month (Weeks 9-12)

During the third lunar month, a great deal of behavioral development occurred. Much of this may be looked upon as an increase in vigor of response and in frequency of response. General activity became greater. The subjects' eyes followed us more persistently than before, they twisted their heads far to one side to see us and they smiled at us more often from week to week. Vocalization also became more common in the course of this month.

Several responses appeared which were relatively new. Cooing, a new vocalization, was heard repeatedly. Laughter also had its inception in this month. Another new pattern which occurred with Del during the third month, but later in Rey's case, was that of strongly raising the head when she was lying on her back. She gave the appearance of trying to rise to sitting, although she had never seen anyone perform this reaction. Del also made apparent attempts to turn from the back to the side position.

Just at the close of the month Rey began to watch her own hands, and on one occasion fixated one hand steadily for 2½ minutes.

Changes of an emotional character were noticed. During the early part of the month, the babies, if they were hungry and crying as we entered the room, would cease their cries after our entrance and remain quiet while waiting to be fed, even though several minutes were required to complete the preparations for feeding. Toward the close of the month, they became more impatient, occasionally resuming their crying if they were not fed soon after our entrance. Crying sometimes started as the bottle in full view of the child was being carried toward her. The twins were particularly likely to cry, if, once the bottle was in view, it was withheld for a few moments.

C. FOURTH LUNAR MONTH (WEEKS 13-16)

Both infants frequently engaged in watching their own hands, although when we were near them their attention was directed toward our faces. During this month, the subjects occasionally cried when we walked away from their cribs. In this circumstance, we did nothing to reward these cries. Such cries were never severe or prolonged.

If we picked up one of the subjects as if to feed her and then returned her to her crib, very hearty crying always ensued. This occurred even though the child had not previously been crying. The cries upon being returned to the crib were much louder than cries due to hunger alone.

The subjects in the fourth month noticed a stranger for the first time. That the twins distinguished between a stranger and a familiar adult was shown by prolonged sober-faced staring at the unfamiliar person whereas when the gaze was directed toward us it was neither so fixed nor so serious.

Motor development was not very marked, but the subjects were more active than they previously had been. Rey attempted to turn from back to abdomen, as Del had attempted it earlier but neither as yet succeeded. In trying to rise to sitting Del not only lifted her head from the bed but also raised her shoulders. Del managed to clasp her hands together.

D. Fifth Lunar Month (Weeks 17-20)

By the seventeenth week, the infants smiled almost invariably when we approached the cribs. With Rey this amiable and untrained greeting went beyond the act of smiling so that she occasionally laughed as we appeared. Several tests of laughter were made in the fifth month. The twins laughed freely to several forms of stimulation.

Attention to sounds became much more marked. Rey in particular began to look about for the source of sounds at this time, a response which was less noticeable in Del. Del, however, was much more afraid of sounds than was Rey. Several times during the month loud noises caused Del to cry. In the 18th week we crumpled some wrapping paper in which laundry had been returned to the nursery. Although this had been done earlier without producing any effect upon the twins, on this occasion Del broke into severe crying which lasted for five minutes. Rey looked in Del's direction but did not cry. Del again cried at the crumpling of paper later in the week. In the following week Del cried at the sound of thunder whereas Rey did not do so. In order to determine whether or not singing would cause Del to cry, a song was sung in the nursery. This was the first occasion on which anyone had sung in the twins' presence. No words were sung; the tune of Santa Lucia was hummed by Mrs. Dennis, who kept her lips closed. Rey merely watched the singer, but on each of six occasions Del frowned and puckered her mouth as if about to cry. During this same week, the 17th, our daughter cried in a neighboring room. Again Del cried at the sound, but Rey did not. We shall have occasion later to comment on the fact that whereas Del cried at noises, she never cried in response to Rey, even though Rey's cries were often quite loud.

In the 18th week, the twins were placed together in a crib for a few minutes. Each regarded the other, but no other behavior ensued.

Several forms of motor development took place in this month. The watching of hands continued, but in addition the hands occasionally clasped the bedclothing with which they came in contact. This often led to a considerable amount of pulling of the bedclothes. The hands also clasped each other. The experimenter's clothes were sometimes grasped. In this month appeared the familiar response of putting the hands to the mouth, both with and without the prehended blanket or sheet. The hands were chewed but they were seldom sucked.

The raising of head and shoulders from the bed became more frequent and more pronounced. In two instances in the 19th week Del performed this response when she was being held in arms for feeding, and nearly succeeded in bringing herself to a sitting position. It was, of course, a part of our plan not to aid such efforts.

Vocalization became more varied and more extensive. On one occasion in the 18th week Del vocalized almost continuously for half an hour. We were not in the nursery but we could hear her from an adjoining room. The cries of each twin became more differentiated, approximating a "coaxing" and a "scolding" tone when the bottle was withheld.

E. Sixth Lunar Month (Weeks 21-24)

During this month the responses which had been established earlier became more regularized, more skilled, and more vigorous. Whereas in previous weeks the bedclothes had been pulled only slightly, now more force was exerted with the result that occasionally the beds were thoroughly disarranged. Likewise vocalization became stronger so that at times the sounds of the twins were veritable shouts. Laughter became more common, sometimes occurring when there seemed to be very little external stimulation to call it forth. General good spirits and an air of well-being characterized the subjects.

Among the new responses of the month were the following: grasping the toes or some part of the foot and carrying the foot to the mouth; holding the legs in the air as the infant lay on her back; and resting the hands on the knees. There began in this month a response which for some time seemed to serve as an expression of interest, of wideawakeness, of excitement and of well-being. This we called "extension of the extremities." In this response the arms or the legs or both would be rigidly extended and perhaps moved slightly. This often happened when we bent over the crib, but it also occurred when we were not near. It was frequently accompanied by a smile or by a loud vocalization.

The twins began to play at sticking out the tongue. We need scarcely repeat that in this respect the twins' behavioral development was without example on our part, nor did the twins see each other while they were engaging in this sport. It may be that irritation from developing teeth had something to do with the persistent protrusion of the tongue.

In the last week of the month Del succeeded in rolling over for the first time. As we were engaged in bathing Rey at the moment that Del managed to do this we failed to witness it, but she undoubtedly did it in the manner in which she had tried it for many weeks; i.e., by rotating her head and shoulders. After getting herself turned to the prone position she held up her head for 30 seconds, then dropped it and began to cry and then to sob. This may have been due in part to the fatigue caused by her efforts and in part to the novelty of the situation.

On one occasion in Week 24 the twins were placed together in the same crib for a few minutes. Del got hold of one of Rey's hands and carried it to her mouth, as she was accustomed to do with all objects. After a few seconds Rey began to squirm and fret. It will be recalled that during routine care the twins could see each other only when being fed. They seldom paid any attention to each other at that time, as they were more interested in the bottle and in the experimenter who held it.

It was during the sixth lunar month that the twins were transferred from the room in which we first placed them to a similar room in a summer cabin 17 miles away. The twins' faces were covered during the trip, which was made by auto. The journey seemed to disturb them in no way, nor did their transfer to a new room cause any upset in their usual calm behavior.

F. SEVENTH LUNAR MONTH (WEEKS 25-28)

The seventh month was relatively uneventful. The twins watched us attentively when we were in the room. They smiled and laughed when we came near. Del laughed at the slightest touch from us, often laughing even when we pulled a diaper from under her buttocks or when we removed her clothes for the bath. They showed a slight increase in the tendency to fret or cry when we turned away.

The most common activities of the infants were: extending the extremities, engaging in random limb movements, clasping the hands together, watching the hands, grasping the feet, pulling at sheets and blankets (which we now tucked under to the full width of the mattress), putting hand, foot or blanket in the mouth, chewing the gums, sticking out the tongue, vocalizing, and kicking. The vocalizations became more varied and more conversational in tone.

One of the new responses of this month was the scratching of the sheets and of the wire netting of the cribs with the fingernails.

Del was as easily upset by sounds as she had been formerly. Twice the shaking of rubber sheets beyond her field of vision caused her to cry.

Thinking that the response to sudden and unexpected sounds might be the initial basis for reaction to the scoldings of adults, one of the experimenters in Week 27 said *"ä-ä"* fairly loudly three times at intervals of a few seconds. Del started crying after the first *"ä-ä,"* and Rey began after the third. Scolding was not tried again until the last month of the experiment.

G. EIGHTH LUNAR MONTH (WEEKS 29-32)

So far as new behavior was concerned, this period revealed but little progress. The predominant activities continued to be those of the preceding weeks. One of the new responses was turning from supine to prone on the part of Rey. However, before she accomplished this she had been placed somewhat out of the center of her bed, so that she was on a slight incline. This must have been of considerable aid to her, for she did not repeat the performance for several months. Another new item was the seizure by Rey of the experimenter's hair and face as he leaned over the crib. Del succeeded in grasping the nursing bottle on two occasions. Del's other novel reaction of the month was the development of an affected cough—a cough which she often gave when it was obvious that there was no need for coughing.

There were certain changes in the experimental conditions which were introduced at this time. One modification of the earlier rigid routine consisted in permitting ourselves to talk to the twins and to play with them to a slight extent. Since they had developed very decided positive reactions to us in the absence of such attentions, our questions in that quarter were answered and there seemed to be no need for continuing an utterly impersonal demeanor. Our speech to the twins, however, was limited to casual everyday remarks and did not include baby-talk nor did it include any attempts to encourage the twins to imitate us.

A second alteration in the previous routine consisted in placing the subjects for a short time each day on the abdomen on a quilted pad on the floor. This procedure, which was begun in the 30th week, was introduced because we felt it advisable to give the subjects an opportunity for types of exercise which could not be gained when they lay supine. The infants were observed continuously when they were on the quilted pads.

Lastly, a number of semi-solid foods were added to the subjects' diet at this time. We cannot enter here upon a detailed account of the reaction of the twins to new foods.

H. NINTH LUNAR MONTH (WEEKS 33-36)

During this period the subjects were returned to the original nursery in Charlottesville. The journey was without incident, the babies being quiet throughout.

Soon after the return trip, the subjects were taken to the hospital for a routine medical and anthropometric examination such as they had not had during the summer. Both children were very sober in the doctor's office. They looked about, stared, and Del fretted slightly. During the taking of the measurements, Rey cried.

We continued placing the subjects on pads on the floor for a short period each day. Rey learned to progress slightly when prone by bending from side to side. Del developed the ability to turn from prone to supine, and since she could already execute the reverse movement, she did a certain amount of rolling. In the course of these progressions, each infant bumped her head against the wall. When Rey did this she cried but briefly, while Del's accident led to several minutes of sobbing and of holding of the breath.

In the course of being fed some cereal which she did not want, Rey once cried and threw herself backward away from it. Formerly she had either let it run from her mouth or had refused to open her mouth.

I. TENTH LUNAR MONTH (WEEKS 37-40)

In this month we conducted the tests for sitting which are described in detail elsewhere (2). It is sufficient here to state that the twins, who had reached the age of 263 days without being placed in a sitting position, could not maintain a sitting position unaided when they were first placed in it. They fell in any direction in which they leaned. When they fell forward, they supported themselves on their hands as they were accustomed to do when they were placed prone. After this fact was determined, the infants were repeatedly placed in a sitting position and inclined forward so as to sit supporting themselves with the hands. Practice in this position, without instruction or reward, led to a rapid improvement in sitting.

Otherwise the responses of this month were largely the same as those described earlier. There were a few new items, as shown by Table 1. One of these was Rey's attainment of turning from prone to supine.

J. ELEVENTH LUNAR MONTH (WEEKS 41-44)

In this month we often placed the twins in sitting position in their cribs. Neither could as yet get herself to a sitting position. Del often raised her head and shoulders, as she had done several months earlier. In this month she added to this pattern a certain amount of pushing with her elbows, but this did not result in raising herself to the sitting position.

Rey had practically ceased raising her head and shoulders when supine as if to sit. In Week 42 we placed our fingers in her hands and she held on as we pulled her to the sitting position. This Rey often repeated thereafter, but Del would not hold on with sufficient force to be raised in this manner.

Tests for reaching for a dangling ring which took place in this month are described in an earlier report (2). The infants had grasped objects prior to this time, but the objects had known positions (the objects were the toes, the blankets, and the bottle), so that the manual activity may not have been visually directed. The dangling ring was not an object of known position. It was reached for by each subject only after it had been presented many times, and on the first attempt the reaching was not accurate.

At this time, the placing of the infants upon pads on the floor was discontinued. The newly-established sitting position provided a change of posture, and also an additional form of exercise.

The twins were placed in high-chairs for a short time daily beginning with the 43rd week. Soon thereafter they were fed in the chairs. At the start, neither subject liked the novel position. Each had some difficulty in maintaining herself in a comfortable attitude, tending to relax and scoot down in the chair and then being unable to extricate herself from the difficulty. This often led to crying. Proficiency in use of the chairs increased gradually.

K. TWELFTH LUNAR MONTH (WEEKS 45-48)

The twins became accustomed to the high chairs, and spent much time in them.

In order to get measures of handedness, we presented to the twins during this and later months many test objects. Some objects were placed on the chair tray equidistant from the two hands. In other cases we merely observed the play behavior of the subjects in order to see which hand reached for the feet and which hand was put into the mouth. A further test of handedness consisted in removing the nipple of the bottle from the mouth and noting which hand was used to replace it. These tests are described in full in a separate report (3).

Del was found to be generally right handed, so much so that it was with difficulty that she could be brought to use her left hand when her preferred hand was restrained. Rey's handedness was much less simple. In some activities she was right handed, for other actions she used the left hand, and in still different situations she used the two hands interchangeably. It should be remarked, furthermore, that movements in which a certain hand was preferred at one period, occasionally came to be executed by the opposite hand at a later date.

In the course of the twelfth month several new vocalizations were developed. One of these, which was practiced by both infants, was a clucking sound, or click made by placing the tip of the tongue against the fore part of the mouth and then forcibly withdrawing the tongue. This sound was often made when the twins were watching us. Another sound was produced by blowing between the loose lips, causing them to vibrate. Finally, on the last day of the month, Del said *"da-da"* for the first time, although it was, of course, not associated with any object or person.

During this period we took a long series of movie records of the various modes of crying of the infants. Our aim was to obtain moving pictures which would enable us to examine the general bodily expression of the twins when they were crying from hunger, from being restrained, from being picked up and put down without receiving further attention, when they were crying because of loud sounds, and from a variety of other causes. Our technique was to place the movie camera above the subject's crib, stimulate the infant (in the case of hunger, this meant letting the infant go unfed beyond the usual feeding time), and then turn on the camera and photograph the crying baby. The room was brightly illuminated by floodlights for some minutes in advance of the adjustment of the camera.

The results furnish an instance of how readily research plans go astray because of some unforeseen development. The factor which we had failed to take into account in this case was the subjects' strong reactions to ourselves. Although the infant might be crying very loudly from hunger when we approached her, as soon as we reached the side of the crib, the cries ceased. If we stood immobile for several minutes, the fretting might begin again, but the slightest motion on our part, such as reaching toward the camera, was sufficient to inhibit the child's reaction for a further period. When we restrained the subject by holding the infant's head or arms or legs motionless by means of the firm pressure of our hands, the child gazed into our faces and *smiled*. Sometimes the smiling persisted throughout five minutes of forceful restraint. Eventually, fretting would begin, although the experimenter was often fatigued before the subject began to cry. Then, as we noted above, the action of the experimenter in moving his hand toward the camera often put an end to the protest which had just commenced.

The attempt to photograph negative emotions showed that the positive response to persons was strong enough to overcome most instances of negative response. We did, however, succeed in obtaining pictures of some cries. One stimulus which was effective in eliciting loud and prolonged crying despite the presence of the experimenters was a loud noise, made by sounding an automobile horn in the room. Del, in particular, was strongly affected by this stimulation.

L. THIRTEENTH LUNAR MONTH (END OF FIRST YEAR, WEEKS 49-52)

Vocalization was frequent, and consisted very largely of duplicated syllables. Among the sounds made by one or both infants in this month were *baba, dada, lala, and byby*. In Rey's case, there was some evidence that she spontaneously formed an association between the writer and the utterance *"dada,"* although this connection was far from invariable. As stated earlier, we had avoided speaking baby-words in the nursery.

In the 49th week, the first toy was presented to the twins. Each subject was given a rattle, which at first we placed on the tray of the high-chair. A full description of the resultant behavior must be reserved for special presentation, but we may summarize the matter briefly by the statement that Rey from the start was favorable to the sound which was made by moving the rattle, whereas Del in the beginning was frightened by it. Within a few days, however, Del as well as Rey was interested in making noise by means of the toy, and the method of making noise which was hit upon by each twin was that of shaking the rattle, although such an action had never been demonstrated to them. This motion was, however, merely an application in a new situation of the already-established pattern of patting and beating. As soon as each child became active with her rattle, she fretted and cried when it was lost or when it was taken from her.

At almost any time it was possible for us to cause the twins to laugh. On many occasions it was sufficient for one of us merely to appear at the crib and say *"hello."* Shaking our head from side to side, or jerking it toward the infant would also cause laughter. Del once laughed repeatedly when the edge of a blanket was swung against her face. At this period Del began to laugh when her ribs or chest were tickled, whereas several months earlier she had reacted negatively to these stimuli.

In spite of the general good humor of the subjects it was quite possible to cause them to cry so long as we were not attentive to them. We have already mentioned the crying which occurred when the rattles were taken away or when these toys dropped from the tray of the high chair. Rey once cried when a blanket, with which she had been playing, was removed from her crib.

The subjects were negative also to situations which involved elements of unfamiliarity. Strange persons who entered the nursery at this time as well as somewhat earlier, were met with silence and soberness and in Del's case by puckering of the lips. Our normal appearance induced vocalization, smiles and laughter, but we discovered that if our appearance was made unusual we were greeted as if we were strangers. On one occasion during this month one of us donned a paper dunce-cap and went into the nursery. Both children stared with wide-open eyes and Del came near to crying aloud.

Beginning in the 52nd week, the twins were again placed on pads on the floor. Although 10 weeks had intervened since they had last been placed in this situation, no new behavior was elicited. Del rolled, as she had previously done. Rey remained on her abdomen, pulled herself this way and that with her arms, and by this means locomoted to a slight extent.

In the crib situation, the babies developed a new motor item. Both subjects, seemingly in lieu of real sitting, would, when lying on the back, raise the shoulders and chest forward and support herself in this semi-sitting position by resting on the elbows. Neither child could yet raise herself to sitting unaided, although Rey could pull herself up if she was given the hand of the experimenter.

On the last days of the first year, we tested the twins for the ability to support the weight of the body on the legs. Up to this time the children had been limited, so far as bodily posture is concerned, to lying supine, to lying prone, and to sitting. We had carefully kept them off of their feet. In the ordinary home, the mother often holds the young child against her chest, which enables the child to push with his feet upon the lap of the adult, or the mother grasps the infant by the chest and holds the baby so that his feet touch the floor. These common reactions we had refrained from as we wished to know whether the child of one year would support himself in an upright position if he had had no practice in this situation.

Since all children with normal training will support the weight of the body in this position by 12 months of age, a test made at the end of the first year might be said to be decisive as to whether practice has anything to do with the establishment of body support.

The test showed (see 2) that on the first trials on Day 364 the twins made no real efforts to stand, simply giving a few momentary pushes with the legs. In the course of many repetitions of the situation, these pushes became more effective, and before the close of the day of the first test, the twins were supporting their weight for several seconds. By the end of four days of practice, the subjects were able to support their weight for several minutes when they were balanced by the experimenter.

N. FOURTEENTH LUNAR MONTH (WEEKS 53-56)

Imitation, both verbal and manual, was elicited, but the only acts which were imitated were those which were already in the twins' *repertoire*.

New reduplicated syllables appeared, and some of these were relatively specific to certain situations. The twins almost always vocalized when we entered the nursery.

Rey succeeded in raising herself to sitting without aid of any sort. She also stood holding furniture when placed in that position. Del at this time was very much retarded in motor development in comparison with Rey and also in comparison with standard norms.

O. FIFTEENTH LUNAR MONTH (WEEKS 57-61)

In the final month of the experiment both babies made great progress in the field of motor development, but Del's achievements were fewer than were Rey's. Rey pulled herself to a kneeling position in the crib, walked on her knees while holding the side of the crib, walked when led, got herself into a creeping position, and finally crept in her crib. These coördinations were not merely performed once, but in each case the reaction became an habitual one.

In contrast to Rey's marked progress, Del's new acquisitions were few. However, her behavior repertoire was by no means static; it underwent several noteworthy changes in the course of the month, for during this time Del added to her capabilities the power to stand holding furniture, to rock herself on her right elbow when in the semi-sitting position, to pull herself to sitting by holding our hands, and to get herself to the sitting position unaided by rising on her right elbow and then giving herself a final push with the right forearm. But these acquisitions should not obscure the fact that Del was far behind the motor norms for children of 15 months.

Del's retardation could not be said to have been due to a lack of opportunities for learning. Early in this month we decided that we would return the twins to their mother at the end of the month. After reaching this decision we altered our previous rules and gave Del a considerable amount of special attention in an attempt to bring her up to the normal performances for her age. We placed her in the creeping posture and tried to get her to make attempts at creeping. She always dropped immediately to the prone position. We often placed her on a pad on the floor and put a rattle of which she was fond beyond her reach, with the aim of stimulating some trials at creeping. She was fairly adept at rolling and usually managed to secure the lure by means of this form of progression; she did not creep. We led her many times in an attempt at getting her to walk, but she would not try. We often stood her up in the crib, in the hope that this would develop an inclination for standing, for pulling to standing, and eventually for walking. She could stand in the crib when placed against the side of it, but she preferred to sit down or to allow herself to slump and fall. She was not weak, for she expended considerable energy in turning and rolling in her bed, and she bounced vigorously on her elbow when in the semi-sitting posture. She was healthy, and was most amiable. But our efforts at training were apparently without effect and she was unable to creep or to pull to standing at the end of the 15th lunar month (the 14th calendar month).

We placed the infants together many times during this month, and kept a systematic account of the development of inter-twin behavior. To trace the development of interstimulation of the twins we must begin with an observation which was made in the 14th month. The twins had been placed prone on the floor in close proximity in the 52nd week. They had been attentive to each other, and had reacted to each other more persistently than to anything else. Rey rested her head against Del and grasped Del's head, face, and hair. Del vocalized a great deal while this was occurring, and in turn took hold of Rey's hand. After 30 minutes, they gradually lost interest in each other.

The next occasion on which the twins were put together is memorable because it was on this date that real romping developed. This was in the 57th week. The subjects were placed in a sitting position in one of the cribs. Del patted her own knee and vocalized. These reactions had often occurred when she was in her own crib. Formerly they had not elicited any particular attention from Rey. Now, however, Rey laughed each time Del patted her knee, and this seemed to encourage Del to persist. Rey laughed at least 25 times in response to the behavior of her sister. Rey then bent over toward

Del, and pushed her face against Del's ribs. This brought peals of laughter from Del. Del's laughs induced Rey to laugh also, and in addition they caused Rey repeatedly to push against Del, so that the romp, which was very hilarious, continued for some time. It was 20 minutes from the time that the twins were placed together that the romp came to a stop. The end seemed to be caused by fatigue, and by loss of interest in each other.

In interpreting this romp, it must be borne in mind that while the twins had never before stimulated each other in this manner, that each had been shaken and tickled by us. The responses of the twins to each other on this occasion were similar to their responses to us when we romped with them or when we tested various stimuli for laughter.

When the twins were placed together following the incident described above they tended to stimulate each other in ways similar to those which occurred in the course of the first romp. They vocalized to each other, grasped parts of each other's bodies, touched heads together or touched the head to the other's body, inserted the index finger into the other's mouth, etc. These actions often caused laughter, but seldom was it as loud or as prolonged as in the first instance. On one occasion in Week 60 Rey caused Del to topple over, and then lunged upon her. Del laughed loudly and heartily at Rey's treatment. On another occasion, Rey was placed on the floor outside of her sister's crib. When Rey patted the crib, Del laughed. Again, upon being placed in one crib, Rey grasped Del's foot and carried it to her mouth. This also was amusing to Del. When they were together for a period of time, they became indifferent to each other as they had at the time of the first romp. Never, however, did they show aggression, or cause each other to cry. Once Rey attempted to explore Del's eyes with her forefinger, but Del avoided injury by turning her head. It must be borne in mind, in this connection, that these periods of interstimulation were never permitted to continue for more than 30 minutes.

Of vocalization in this month it may be said that it showed progress, but that neither child came to use any sound as a real word within the period covered by the experiment. There was no attempt on our part to train the subjects to use words. However, we often tested them for their willingness to imitate us when we uttered sounds which had previously been made by the twins. In many instances, they imitated us readily under these circumstances.

We took occasion during the final month to determine the twins' reactions to many objects which, because of their limited environment, they had never seen before. In general, the tests reveal the fact that the twins were attentive to new stimuli but these novel objects as a rule caused the subjects to be at first silent and motionless. Seldom could it be said that they showed fear, although they might have done so if their introduction to the objects had been abrupt and sharp. It was characteristic of the twins that after a period of silent regard, they began to make contacts with the test objects. For instance, on one occasion both subjects stroked a large English Pitt bulldog within a few minutes of the time that he was brought into the room.

By the close of the experiment the twins were relatively less affected by strangers than they had been during earlier months. While they were usually of serious countenance during the first few minutes of a visit by a stranger, this manner gave way to smiling and to vocalization within a short time, and it was sometimes followed by actual approach toward the stranger.

The twins' responses to us continued to be of the most positive character. Almost invariably they smiled and vocalized as we entered the room. On the other hand, they did not protest when we left them as they had done earlier, and they practically never cried when we were out of the room. In the opinion of others—and in our opinion, also—the twins were model babies, healthy, happy, active, and completely unspoiled. There was no behavior, save Del's inferior motor development, which distinguished them from other children of the same age.

IV. A LIST OF DEVELOPMENTAL ITEMS

A description such as that which we have just presented cannot propose to name all of the items of behavior which appeared in the course of the experiment. It seems desirable to present a complete list of new responses which made their appearance in the twin's repertoire, and to accomplish this purpose is the aim of the present chapter.

The items which follow are arranged in the chronological order in which they occurred in the experiment. That is, each item was added to the list as it occurred, regardless of the twin which first performed it, although of course the records of the subjects are separated. The two columns on the right side of the table show the age at which the response occurred. If each subject performed the response, the date of appearance is given for each subject.

This table (Table I) is limited to items which occurred *spontaneously*. This means that the table contains responses which, so far as the circumstances of the twins were concerned, *could* have occurred at any time. We have not placed in the list any item which required special action on our part—such as sitting when placed sitting, or supporting the body weight when held upright. Such test behavior *was not included, because, since the tests were not made daily or* even weekly, the period at which test behavior appeared was in large measure a function of the time at which the test was given.

The list is defective in regard to vocalizations. To have listed each distinct sound as it was produced would have extended the list beyond reasonable proportions.

In commenting upon this series of developmental items, we wish to call attention first of all to the extensive character of the list. One hundred and fifty-four distinct responses are listed. The list is by no means complete, for vocalizations were largely omitted. The record shows how marked is behavioral development under conditions of minimum social stimulation and restricted practice.

It will be seen that the order of appearance of the items was roughly the same for the two subjects. If a rank-order correlation is computed for the items which were common to the repertoires of the two subjects, it is found to be .99.

The list shows that during the early months there was little difference in the rate at which new responses were acquired by the two subjects. Del was slightly in advance of Rey. In later months, how-

TABLE 1

NEW ITEMS OF BEHAVIOR

No.	Behavior items	Week of first appearance	
		Rey	Del
1	Hands close on objects	7	6
2	Smile	6	6
3	Cry tears	8	6
4	Fixate *E*	6	7
5	Smile when fixating *E*	6	7
6	Frown	6	6
7	Turn from back to side	8	6
8	Tense during bath	6	
9	Arms active in bath	9	7
10	Cry at sound	18	7
11	Corners of mouth pulled down before crying	7	7
12	Vocalize vowel	7	8
13	Rub face	9	7
14	Arm startle to touch or jar	7	7
15	Tongue between lips	7	7
16	Turn head to watch *E*	8	7
17	Stop crying when *E* enters nursery	7	8
18	Stop crying when light turned on (at night)	8	7
19	Open mouth smile	11	7
20	Scolding vocalization	8	17
21	Fret	8	8
22	Coo	8	8
23	Hand to mouth	8	10
24	Vocalize when fixating *E*	9	9
25	Shake head	9	11
26	Purse lips	12	9
27	Raise eyebrows	9	9
28	Fixate objects beyond crib	11	9
29	Jerk upon hearing noise	11	9
30	Consonant-vowel syllable	9	
31	Near-laugh	10	9
32	Eyes watch *E*'s hands	10	9
33	Noise inhibits crying	10	10
34	Look for *E* when he disappears behind screen	10	10
35	Vocalize single consonant	10	36
36	Stick out tongue	10	25
37	Cry when *E* leaves crib	18	10
38	Rub eyes	10	10
39	Lift head when supine	16	10
40	Cry upon seeing bottle	11	10
41	Fixate own hand	12	11
42	Feeding inhibited by staring at E	11	26
43	Cry when put down without attention	15	11
44	Chewing movements	12	12
45	Crowing vocalization	12	12
46	Attempt to turn from supine to prone	16	12

TABLE 1 (*continued*)

No.	Behavior items	Week of first appearance Rey	 Del
47	Turn toward sound	19	12
48	Carry object to mouth	21	12
49	Increased activity at sight of *E*	15	13
50	Watch own hands persistently	14	14
51	Kick when otherwise quiet	15	14
52	Coaxing vocalization	14	20
53	Stare at stranger	14	14
54	Stop nursing to coo or smile	17	16
55	Sob		16
56	Clasp own hands	18	16
57	Laugh	21	16
58	Gurgle	16	16
59	Raise head and shoulders when supine	18	17
60	Turn supine to prone	50	17
61	Head inclined toward approaching bottle	19	18
62	Pull at bed clothes	19	18
63	Open mouth when *E*'s finger approaches face	20	19
64	Arms extended, expressive movement	19	22
65	Pucker face at stranger		20
66	Place hands on knees	20	20
67	Smack lips when being fed		20
68	Hold legs in air	21	21
69	Grasp own ankle	22	21
70	Rising vocalization when nipple removed	22	22
71	Extend and wave arms	27	22
72	Grasp own toes	22	22
73	Laugh during solitary play	23	23
74	Place soles of feet together	23	
75	Shouting vocalization	27	24
76	Grasp own diaper	24	24
77	Cry when *E* leaves room	25	24
78	Wiggle toes	27	
79	Scratch sheet or crib	27	29
80	Open and close hands repeatedly	27	
81	Touch hands together	27	27
82	One hand clasps other hand	30	27
83	Sing-song vocalization	29	27
84	Grasp own ear	31	28
85	Grasp own hair	29	
86	Scratch own head	29	
87	Affected cough	30	29
88	Pincer prehension of bedclothes	30	29
89	Fret when bib put on	31	30
90	Arch back when supine	30	31
91	Grasp *E*'s face and hair	31	
92	Pat own body	31	33
93	Pincer prehension of own toes	31	

TABLE 1 (*continued*)

		Week of first appearance	
No.	Behavior items	Rey	Del
94	Open and close mouth repeatedly	31	
95	Kick during bath when otherwise quiet	33	31
96	Suck lower lip	32	31
97	Turn prone to supine	39	33
98	*Roll on floor*	51	33
99	Pat object	34	45
100	Throw self backward in rejecting food	34	
101	Pat own foot	34	36
102	Smack lips when not eating	35	37
103	Cross legs	35	35
104	*Pat knee with contralateral foot*	36	
105	Rock self on abdomen		36
106	Rest head on palm of hand when supine	37	49
107	Rub soles of feet together	37	
108	Vibrate lips by blowing	44	37
109	Toe to mouth	37	41
110	*Pivot on abdomen*	43	37
111	Vocalize click	37	49
112	Pat own hands		37
113	Monotonous vocalization while going to sleep		38
114	Foot push against crib	38	
115	Suck at sight of bottle		39
116	Reach for spoon when being fed	39	45
117	Snap fingers	41	
118	Bend forward and rest on elbows when supine		41
119	Go from sitting to prone position	46	43
120	Put covers on and off face	44	
121	Rubs own gums		47
122	Grasp toes while sitting	47	47
123	Bounce while sitting	48	49
124	Cry when object taken from her	49	48
125	Touch and pursue object with forefinger		49
126	Push floor with foot when prone	49	51
127	Display repertoire for stranger		49
128	Scoot backward when prone	51	
129	Splash in bath	52	54
130	Whisper	52	54
131	Rise to half-sitting, resting on elbow	52	52
132	Take nipple from mouth, plays with it		54
133	Finger lips while vocalizing		54
134	Rub self in bath	55	
135	Raise self to sitting position	56	61
136	Reach in diaper for feces	57	
137	Cross index and middle fingers	57	
138	Pull to kneeling position	57	
139	Kneel without hand support	58	
140	Shake bedclothing		58

TABLE 1 (*continued*)

No.	Behavior items	Week of first appearance Rey	Del
141	Walk on knees, holding side of crib	58	
142	Shrug shoulders		59
143	Pull to standing in crib	59	
144	Go from standing to kneeling in crib	59	
145	Put foot on tray of high chair		59
146	Rock self on elbow in half-sitting position		59
147	Scoot backward when sitting	59	
148	Rock when sitting		59
149	Get to kneeling position without pulling	60	
150	Progress by bouncing in sitting position	60	
151	Pivot while sitting	61	
152	Stand momentarily without support	61	
153	Creep	61	
154	Walk, holding to crib	61	

ever, Rey was considerably more advanced than Del, especially with regard to locomotor items.

Finally, we may note that behavior became more individual in the later months. At the beginning, the subjects almost invariably developed the same responses. In the later months, each subject developed many items which were never practiced by the other twin.

V. DISCUSSION

A. AUTOGENOUS BEHAVIOR

Our chief interest in conducting the experiment lay in determining the responses which a child would develop if he were removed as far as is possible from the influence of adults and of other children. This is an old query, having been asked several centuries B.C. by Psametichus who reputedly reared two children apart from speaking adults in order to determine whether or not they would develop language in the absence of adult example (8). The question as to what the unaided individual would be able to do was the basis also of interest in the so-called "wild children" who have been found from time to time, the most famous of which was Victor, the wild boy of Aveyron, who was studied and trained by Itard (9). Since there has been a prolonged interest in knowing what responses the individual is able to achieve when uninstructed, it is surprising that no term by means of which we can designate such responses has come into general use in psychology.

The term "instinctive response" immediately comes to mind. To be sure, this term has been accorded a variety of meanings so that it would not be difficult to find one usage which approximates our present need. But there are many objections to the use of "instinct" in this sense. There is a tendency at the present time to limit the term "instinctive response" to the connotation of a reaction pattern which requires no learning, and it seems best to further the acceptance of this usage. We wish a term which will not only include "unlearned response" as defined above but which will also include other responses which are developed by the individual independently of other persons, whether by discovery, by trial and error, or by whatever means. We wish a term which will designate instinctive responses and in addition responses which are independently acquired.

We suggest that such responses be called "autogenous."[2] This word already has a recognized place in the vocabulary of several sciences, and is capable of useful application in describing behavioral development. The general meaning given for autogenous in Webster's *New International Dictionary* is "self-generated; produced independently of external aid; endogenous." In medicine the term is used to refer to some condition which originates within the organism as opposed to a condition which results from contagion. In anatomy, autogenous is employed to refer to ossification which proceeds from an independent center. In zoölogical usage, the term refers to a condition which originates within the organism. These meanings will be seen to be in line with the one suggested above for the description of behavioral development.

---

[2] This suggestion was first made in a paper presented at the Annual Meeting of the American Psychological Association in 1936 (5).

The present study has shown to what an extent autogenous behavior is characteristic of infancy. While a child cannot be excluded from all contacts with other persons, we feel that in the present experiment our contacts with the subjects did not provide the *origin* of the new responses which were recorded. We avoided the use of reward and punishment. We refrained from performing the infantile responses which the twins added to their repertoire. From our experiment it appears that practically all of the behavior of the first year of life is autogenous. The diary account of development gives an impression of behavior as extensive, as varied, and as typically human as does a biography of an infant reared under normal home conditions. In a previous report (6) it was shown that the twins exhibited all of the responses which are commonly recorded by the biographical method, and also all of those which Shirley found to be characteristic of her subjects. The extensiveness of the autogenous repertoire is indicated by the fact that we were able to list 154 new responses which occurred in the course of the experiment.

No doubt the introduction of training might have increased slightly the number of new responses. The additional reactions would have been such items as waving the hands when the adult said *"bye-bye,"* of kissing and of throwing kisses, and of pointing to objects when they were named. Such responses are obviously sociogenous, and have been considered so for centuries because they never appear until after appropriate training has been administered. But it is likely that the number of sociogenous responses which can be acquired by a child under one year of age is very small. They are in no sense an essential part of normal development during the first year, and since they are unusual at that period they do not constitute an exception to the statement that our subjects autogenously exhibited all of the common responses which are characteristic of infants in the period preceding the first birthday.

B. THE QUESTION OF INSTINCTIVE RESPONSES

How many of the autogenous responses of the first year are instinctive? Before we can answer this question we must determine how one is to discover instinctive responses and we must then canvas our data from this point of view.

As used here, the terms "instinctive response" and "unlearned response" are synonymous. We explicitly exclude from the discussion all other connotations which have been associated with the word "instinct."[3]

An unlearned response is usually defined in negative terms as a reaction which has not been learned. Since we have no universal agreement as to what constitutes learning, and no infallible indication of its contribution to the development of a response, it follows that we do not have an operational method of identifying instinctive responses.

This much, however, is probably agreed upon by all students of behavior, that if it were rendered impossible for an organism to display responses before a given moment—as by administration of an anesthetic—that all the responses which appeared upon the organism's first opportunity to react might be called unlearned responses. In other words, it is commonly agreed that learning cannot occur in *the absence of a reaction history, and hence all responses which appear* coincidentally with, or shortly following, the onset of motility may be considered instinctive. But since it is not reasonable experimentally to render the human fetus or the human infant immobile, this criterion of an instinctive response has no applicability to man.

A second basis on which some agreement might be reached has reference to the *kind* of reaction history which precedes a specified response. We are interested to know whether a response suddenly appears in full or whether its final form is achieved only after a series of approximations. If a response occurs suddenly, without any forerunners which resemble it, we are inclined to think that it may be unlearned. However, we cannot be certain of this interpretation. A response may develop through learning without our being able to recognize this fact. For instance, if typical human laughter occurs in an infant who has never laughed before, it is obvious that laughter has not been achieved by the gradual perfection of responses which could be recognized as attempts to laugh. Nevertheless, laughter may have been attained by virtue of the fact that the child earlier engaged in smiling, in certain laryngeal responses, and in a variety of types of breathing. In the development of laughter, the child might thoroughly practice every element of laughter without our being able to detect by ordinary observation any rôle of practice in the execution of the first complete laugh, since the practice would not result in the production of sound if the elements were practiced separately.

---

[3] We cannot at this point undertake to refer to discussions of the concept of instinct, nor can we refer to the numerous studies of development which have some bearing on the question of post-natal instincts in man. We hope to consider these materials at a later date. To review this body of literature in conjunction with the present report would not be feasible.

Even if for the present we disclaim the ability to detect an instinctive response we may still retain a legitimate interest in the *ontogenetic history* of responses. We may ask which reactions make their appearance with little apparent contribution from the previous reactions of the individual. Such responses would seem the most likely to be truly unlearned.

C. Some Negative Results

In an earlier report (2) we have presented some negative evidence in regard to instinctive post-natal responses. This study dealt with three responses, which under the conditions of the experiment, it was practically impossible for the subjects to practice. They were: reaching for a dangling ring, sitting without support, and bearing the weight of the upright body upon the legs. Each of these reactions was tested for when the subjects were at an age such that all normal children of normal environment exhibit the response. Nevertheless each response was absent in the first few tests. These tests thus yielded negative evidence as to the instinctive character of these responses. This was also the case in a test of the alleged instinctive connection between infant smiling and the sound of the human voice (1).

It is worthy of note that with repeated contacts with the test situations the subjects, without instruction and without any administration of reward or punishment by the experimenters, gradually developed the appropriate responses.

D. Some Positive Evidence

Three responses appeared which seemed not to be derived from the antecedent reactions of the subjects. The first of these responses was laughter. The other two we shall tentatively call "attempting to rise" and "attempting to turn," although in greater strictness they might be labelled "lift head and shoulders when supine" and "rotate head and shoulders when supine." We shall review the history of each of these actions. Most of the observations to which we will refer have already been presented in the biographical account but need to be again examined with reference to their instinctive character.

1. *Laughter*

Rey's first clear-cut laughter occurred in Week 21, while Del's first unquestionable laugh was heard in the 16th week. Since each infant smiled and vocalized long before these dates, it is possible that the exercise of these two functions had an influence upon the development of laughter. However, only twice in the case of each subject was any sound resembling laughter heard prior to the date of the first unmistakable laugh. These sounds, which will be called near-laughs, were Weeks 10 and 13 in Rey's case and in Weeks 9 and 10 in Del's case. The absence of any other record of near-laughter prior to full-fledged laughter is all the more striking because our observations of sounds were not limited to the time which we spent in the babies' room. We could hear the vocalizations of the infants with ease even when we were not in the nursery.

The occurrence of two sounds somewhat similar to but not identical with laughter can hardly be said to be favorable to the view that laughter develops through practice.

2. *"Attempting to Rise"*

The second response which appeared as if without practice was the raising of the head when the child lay on her back. Often this behavior was engaged in persistently as if its full goal had not been attained. From its inception, it had the appearance of a true attempt to attain the sitting position although neither child could well have been acting from any such conscious motive at the time of the first trials. Rey's performance of this response was first observed in the 16th week and Del's in the 10th. The response recurred from time to time and gradually became stronger. By the 18th week Rey was able to raise her shoulders as well as her head above the level of the mattress; Del did this a week earlier than did Rey. Nevertheless it was only in the 61st and the 56th weeks, respectively, that the two infants were able to attain the sitting posture without assistance. It is more than likely that their many attempts to sit were a real factor in the final achievement of this response, but we are at a loss to know how learning could have contributed to the first lifting of the head from the pillow, and to know why the activity was persisted in when it seemed to achieve nothing.

3. *"Attempting to Turn"*

The third response which gave the impression of being unlearned was the first instance of turning vigorously toward one side while supine.

Del on a day in the 12th week worked continuously for 20 minutes in an apparent effort to turn over from the supine to the prone position. Her efforts did not succeed in changing her position and in the end she cried. Rey in Week 16 made a similar but less strenuous and less persistent effort which also was unsuccessful. It did not result in crying. In these performances the infants' shoulders and head were repeatedly rotated to a great degree. There was no suggestion that the subjects were trying to reach anything or to get a better view of anything. Nothing external seemed to control the reactions. While rotations of the head and slight shoulder rotations had occurred before this time, the performances just described were definitely novel in their extent and their persistence. The stimulus to these reactions may have been discomfort engendered by lying continuously in one position, but whatever the stimulus, it, in itself, cannot account for the patterning of the behavior which resulted from it.

These actions, which we have called attempts at turning, were not repeated by either baby—at least not while they were under observation—until several weeks had elapsed. With Rey the second attempt was seen in Week 24, with Del, in Week 22. Del's third known attempt, in Week 24, was a success, but she cried strongly after she had lain in the newly-achieved prone position for only 30 seconds. She did not turn again from supine to prone until Week 28, and this time also the performance was followed by crying. Crying did not recur in this situation thereafter. Her skill in turning over gradually improved, so that by Week 35 she was accomplishing the feat nearly every day.

Rey was never seen attempting to turn over after the failure recorded in Week 24 until Week 35. Her first success was not until the 50th week. Subsequent to this, turning from supine to prone became a frequent response.

E. MEAGERNESS OF THE EVIDENCE FOR POST-NATAL INSTINCTS

No responses other than these three accorded so closely to the more or less classical notion of what an instinctive response should look like. In making this statement, we are referring, of course, only to the responses which were added to the repertoire which the subjects had at the beginning of the experiment. It may well be that many responses which are observable in the *neonate* are independent of the reaction history of the organism.

In view of the large number of developmental items, it is significant that we saw only three responses whose origin did not seem to lie in the slow modification of antecedent activities. The trend of our records is toward ascribing only a very small unlearned element to those additions to the infant's repertoire which occurred in the course of the experiment. We observed only three responses which *appeared suddenly as if from the maturation of some bodily mechanism.* Although these responses seemed to appear "out of the blue," we must admit the possibility that movements previously made may have contributed to the appearance of these responses. It is also possible that these responses may have had precursors before birth, or between birth and the beginnings of the experiment, or when we were not observing the subjects. For these reasons, we cannot be dogmatic with regard to the mode of development of these activities.

All other responses which we studied presented a picture of a gradual acquisition, and hence at least *simulated* learning. It is possible, of course, that growth processes may affect behavior in the same manner as does learning and hence the gradual improvement *of coördinations may give no clue as to the factors which are responsible* for this improvement.

For the sake of simplicity we have spoken as if response patterns are totally instinctive or totally learned. We have no reason to suppose that this is the case. It seems more probable that instinctive elements and learned elements contribute jointly to the formation of a single response. For instance, when our subjects were first placed on their feet, they did not support themselves but they did give an occasional push with the legs. These extensor thrusts in response to pressure on the soles may have been unlearned. In later trials, extension of the legs became more persistent, more regularized, and more energetic so as finally to lead to support of the body. We have said that the response was learned. By this we mean that some learning was undoubtedly involved, since the response did not appear at the usual *age when opportunities for practice had been eliminated.* However, this does not involve a denial that instinctive contributions may also have been present. Instinctive elements by themselves did not lead to "standing with support" but they may have been essential to the formation of this pattern. In any infantile pattern, instinctive bases and learned modifications may be inextricably mixed. Our aim has been to show that in all likelihood instinctive elements by themselves do not constitute normal infant responses.

VI. SUMMARY

In this report the experimental conditions surrounding the development of our two subjects have been presented in some detail. The aim of the experiment was to throw the subjects as much as possible upon their own resources and their own initiative in so far as behavioral development was concerned. The experiment covered a period of 14 lunar months, extending from the end of the first to the end of the 15th lunar months.

Previously published reports upon certain aspects of the investigation have been reviewed. The new material here presented consists in a month-by-month summary of behavioral development and a list of 154 developmental items.

These data, together with those which have been presented earlier, have been interpreted as indicating two general conclusions:

1. The first of these is that practically all of the common responses of the first year of life may be developed autogenously. That is, infants will develop these responses without encouragement or instruction, without reward or example. It follows that prior to the second year of life sociogenous responses, those which are learned through the intercession of other persons, are few and are relatively unimportant. If the well-being of the infant is assured, his behavioral development will take its normal course.

2. The second conclusion is that in the development of the autogenous responses of the *first year learning* plays an important part. The dichotomy of learned and unlearned responses is difficult to employ, but if we are to apply these concepts the evidence indicates that there is little reason to believe that infant responses are made up exclusively of unlearned elements except in the case of very few reactions. However, responses which involve learning may at the same time involve instinctive contributions. As a corollary of this conclusion, it would follow that while maturation is a major factor in infant development its importance lies chiefly in making learning possible. Maturation in and of itself seldom produces new developmental items, but maturation of structures when accompanied by self-directed activity leads to new infant responses.

REFERENCES

1. DENNIS, W. An experimental test of two theories of social smiling in infants. *J. Soc. Psychol.*, 1935, 6, 214-223.
2. ———. The effect of restricted practice upon the reaching, sitting, and standing of two infants. *J. Genet. Psychol.*, 1935, **47**, 17-32.
3. ———. Laterality of function in early infancy under controlled developmental conditions. *Child. Devel.*, 1935, 6, 242-252.
4. ———. A psychologic interpretation of the persistence of the so-called Moro reflex. *Amer. J. Dis. Child.*, 1935, **50**, 888-893.
5. ———. Infant development under minimum social stimulation. *Psychol. Bull.*, 1936, **33**, 750.
6. ———. Infant development under conditions of restricted practice and of minimum social stimulation: A preliminary report. *J. Genet. Psychol.*, 1938, **53**, 149-158.
7. DENNIS, W., & DENNIS, MARSENA G. Behavioral development during the first year as shown by 40 biographies. *Psychol. Rec.*, 1937, **1**, 349-361.
8. HERODOTUS. The History of Herodotus (trans. by G. Rawlinson). New York: Dutton, 1910. (Vol. 1, pp. 110-111.)
9. ITARD, J. The Wild Boy of Aveyron (trans. by G. Humphrey and M. Humphrey). New York: Century, 1932. Pp. 110.
10. SHIRLEY, M. M. The First Two Years. (3 vols.) Minneapolis: Univ. Minn. Press, 1931-1933.

Published as a separate and in *Genetic Psychology Monographs*, 1941, **23**, 191-234.

AN ANALYSIS OF THE MENTAL FACTORS OF VARIOUS AGE GROUPS FROM NINE TO SIXTY*

Benjamin Balinsky

New York University

	Acknowledgments	193
I.	Introduction	195
II.	Description of samplings and tests	203
III.	Presentation of data	209
IV.	Interpretation of the results	219
V.	Summary and conclusions	229
	References	233

---

*Received in the Editorial Office on June 27, 1940, and published at Provincetown, Massachusetts. Copyright by The Journal Press.

ACKNOWLEDGMENTS

This paper is based on a Ph.D. dissertation done at New York University. Many helpful suggestions were received from my sponsoring committee, Dr. Brian E. Tomlinson, *Chairman*, Dr. William P. Sears, Jr., and Dr. Glenn S. Thompson. To Dr. David Wechsler I owe much, both for allowing me access to his files and for the vital discussions on the subject of this paper. And, finally but not at all least, I must pay tribute to my wife, Ruth, who encouraged and aided me in many ways.

Benjamin Balinsky

Consultation Service
National Youth Administration
265 West 14th Street
New York City

I. INTRODUCTION

A. Historical Background

For many centuries, man has been concerned with the problem of understanding his fellow man. The ancient Sophists gave a new direction to human thought when they turned from speculation about cosmology to the consideration of human affairs. They influenced such later philosophical giants as Socrates, who believed that he could obtain knowledge of the "self," and his distinguished pupil Plato, who recognized individual differences in man. Plato went so far as to propose an ideal state, the Republic, where men were to be chosen for special duties, with reference to their ability to perform them. Aristotle continued the emphasis on human matters and "addressed himself to the things of the mind in much the same manner as to anything else in the material order" (11).

Some of the Greek philosophers even considered classifications for the various types of man. Theophrastus proposed the schemata of some 30 types, into which he thought he could fit all individuals; and Hippocrates described all men in terms of the four "humors" which he had devised. Moreover, the Hindus invented a caste system, the rigidity of which is still felt in India.

This desire to classify and to understand man is very strong, and the tendency has persisted through the centuries that have passed. The persistence indicates a definite need for a smaller number of terms in which to describe so complex an animal as man. At present there are many schemes of classification, using such descriptive terms as introvert and extrovert, and pyknic and asthenic, each of which brings to mind a certain kind of behavior.

We are here concerned with the cognitive functions as measured by tests of intelligence. It is well known that there exist innumerable tests purporting to measure this aspect of man. The first tests or objective measures were specifically devised by Galton and Cattell to measure such "mental abilities" as memory, mental imagery, and association. When compared with other ratings, such as school marks, this type of mental testing was found to have certain deficiencies.

It was not until about 1905, when Alfred Binet devised his test of "general" intelligence, that the field of mental testing began its great expansion. Binet concluded, after some experimentation, that no single performance could be used as a just indicator of intelligence.

From this point on, tests of intelligence increased rapidly in number. Many group tests were published, and tests of a "performance" nature were standardized.

During this time, statistical techniques were being formulated for the evaluation of the tests which had been devised. They had their basis in the mathematics of probability, which were first set forth about the year 1600, and were more thoroughly formulated by such mathematicians as Bernouille, De Moivre, Laplace, Gauss, and Quetelet. It fell to Galton to apply the methods of statistics to anthropometry, and he himself added other tools for this work. With the important contributions of Karl Pearson, statistics became a powerful and useful tool for the purposes of analysis.

During the present period of mental testing, a great deal of attention has been paid to the nature of intelligence, and there has been much discussion as to what specific things the tests of intelligence were designed to measure. It was the extremely important proposal by Spearman (22) on the two-factor theory, in 1904, that stimulated a great deal of research in this field. This theory provided an explanation for certain common findings in mental testing, such as positive intercorrelations, and lent succor to the tests of "general" intelligence, like that of Binet. But was the two-factor theory a broad enough means of explanation for other procedures in mental testing? Apparently it was not. Thomson (6), Spearman's severest critic, propounded the theory of group factors, which offered an explanation for the fact of low intercorrelations among parts of a test battery, and which was more suited to serve as a theoretical basis for vocational testing.

The two-factor theory of Spearman fails to explain why the best batteries of tests are those in which the parts show low intercorrelations with each other and high correlations with the entire battery. For, if the entire battery measures a quality to be designated as *G*, it follows that each part of the battery, in order to correlate highly with the entire scale, must be somewhat saturated with *G*. This being so, the intercorrelations among the parts will also have to be high, for these correlations are in proportion to their saturation with *G*.

Thomson's group-factor theory does give an adequate explanation for the low intercorrelations among parts of the tests. He assumes that there are many elementary abilities, and that each test samples a certain range of these. If two tests happen to sample many of the same abilities, they will have certain things in common, and will therefore correlate to a certain degree, depending on the number of common abilities measured. The various parts of the tests may sample small numbers of the same abilities, and thus have low order correlations. But all parts sample some of the entire number of possible abilities, and thus each part of a test battery correlates to some extent *with the total.*

Concerning the theoretical explanations for the possibility of vocational testing, Guilford has written (10):

> In the first place, if Spearman's *s*-factors are really confined to one task alone, as he insists, there is no way of predicting from tests an individual's standing in his *s*-factors for various occupations. It is likely, since many occupational activities have a low correlation with tests that measure *G,* that occupational tasks depend largely upon their own *s*-factors, which could not be measured without asking an applicant to perform the tasks themselves. . . . Vocational guidance would thus be next to impossible without allowing the individual to try out the different alternatives. *On the basis of the sampling theory,* however, both vocational selection and vocational guidance become logically possible.

Spearman has come to accept the existence of factors other than *G.* He and his associates have admitted that in many instances the tables of intercorrelations may include some correlation over and above *G.* Thus they recognized as group factors verbal ability, numerical ability, mechanical ability, and a possible factor of mental speed.

Since Spearman formulated the two-factor theory and devised the method of tetrad differences for discovering the factors, other procedures have been invented. These allowed for the extraction of more than one factor, as Spearman himself admitted the possibility of several factors in test batteries. Kelley (14), Thurstone (25), and Hotelling (12), each developed techniques of factor analysis.

We shall concern ourselves with Thurstone's method of factor analysis, since it is the one chosen for use in this study. This method of analysis allows for the extraction of many factors, which are called weighted group factors, in order to differentiate them from the group factors of Thomson's sampling theory. Whereas Thomson postulated innumerable elementary abilities, Thurstone (26) has postulated a limited number, not yet fixed. The Thurstone multifactor theory, like Thomson's sampling theory, serves as an explanation for low intertest correlations, as well as high test correlations with the entire battery, and also as a basis for vocational guidance.

Thurstone's centroid method of factor analysis, rather than any other method, has been chosen for study here, because it is the most widely accepted and the most thoroughly rationalized. Alexander (1) selected this method for use after comparing several techniques, and wrote concerning his choice:

> For practical purposes it is our opinion that the method of Thurstone is to be preferred because of its greater simplicity and speed in application. For that reason we propose to adopt this method throughout the work that follows.

Morris (18) also employed the techniques, stating that "Thurstone's center of gravity technique is the general case of which all others are but special cases." In a critical comparison of the Thurstone technique with the Hotelling method, McCloy, Metheny and Knott (17) found the Thurstone method superior.

Wilson and Worcester (29), in a discussion of factor techniques, justify the Thurstone method. They have written:

> The Hotelling analysis has the arithmetic advantages of determinacy and of ready mechanical applicability. These advantages may bring in with them very real disadvantages. There is perhaps nothing more likely to convince one that he has something of value than the ability to execute a mechanical arithmetical procedure. For that reason such considerations as have been offered in this "Note" will not be without usefulness if they seem once more to direct attention to the need of examining a formal solution for its psychological substance. It would be a pity if Wundt had taken Psychology away from her mother, Philosophy, and married her to Science only to have her desert to a paramour, Mathematics.

Thurstone's method of factor analysis, like all the others, starts from the intercorrelations of the various tests included in the battery. These tests may be parts of a single scale, or they may be tests selected from whatever source. When given to the same subjects, the intercorrelations of every possible pair of tests can be calculated. As happens invariably, the correlations between the pairs of tests are almost all positive, indicating some thing or things common among them.

Thurstone has devised a rational method of obtaining the number of elements common to the tests, and then identifying and naming them. The common elements are called factors or abilities, and are extracted from the original correlations until the residuals of the correlations no longer can give up significant factors. A fuller discussion on the matter of how many factors to extract is reserved for a later section of this investigation, where the problem will actually arise.

These factors and their loadings or weights have no psychological meaning. They merely represent correlations between any test and one of the reference axes, and may be positive or negative. The reference axes are the usual x and y axes of a coördinate system. Any test has projections on the two axes. The correlation between any two tests is the sum of the products of the factor loadings of each test, i.e., $r_{ab} = a_1b_1 + a_2b_2 + \ldots a_nb_n$.

To make the factors psychologically meaningful, the reference axes must be so rotated that they pass through or nearby the majority of the loadings. This may be done as often as is necessary, each time with the newly found rotated loadings, until the number of zero loadings is maximum and the number of negative loadings minimum. In this way certain tests will have significant positive loadings in certain factors, and others will have insignificant positive loadings.[1] By studying the tests with significant loadings in the various factors, one may arrive at the elements common to the various test groupings and thus name the factors. An examination of the tests with insignificant loadings will aid in the naming of the factors, by indicating points which are not included in these tests, but are included in the others. Since the rotated reference axes are perpendicular to each other, or orthogonal, the factors or abilities are independent of each other.

B. THE PROBLEM

This study undertakes to analyze the mental factors appearing over a widespread period of years, in order to note any significant changes which may occur during periods of growth and decline. These changes may occur in the factors themselves, so that some will appear and others disappear at different periods of time; or the tests that fall in with certain factors at one time may enter in with different groups at another time. A dynamic picture of mental organization may also be obtained from this study.

---

[1] It may happen that after several rotations a negative loading will still be found. If it is within the limit of —.10, it may be considered as insignificantly negative (10, p. 507).

The opportunity to carry out such an investigation as the present one is afforded by the standardization of the Wechsler-Bellevue *Intelligence Scale* (28). This is a point scale which can be used over such a wide span of ages as from nine years to sixty.[2] By means of this research, it will be possible: (*a*) to isolate the mental abilities at each age level; (*b*) to note their stability from age to age; (*c*) to study the tests as they are accounted for by each factor at each age period; and (*d*) to obtain certain facts as to mental organization.

Thurstone (27), in a description of what he calls the primary abilities, recognizes the need for studying various age groups. He gives as an illustration the solution of the simple arithmetical problem 3 x 1¼. He states:

> To answer a score of such items quickly at the age of fifteen is indicative of the factor *N*, but a four-year-old who solves such items rationally may reveal logical faculties, perhaps inductive, rather than superiority in the factor *N*. This is an example of what might be taking place.

He goes on to point out that "these relations have not been adequately recognized in recent studies of the changes in mental organization with age." At any rate it is "advisable to isolate the mental abilities separately at each age level and to move with caution in extrapolating the factorial interpretation of a test for widely different ages." Kelley (15), in a recent critique of factor analysis studies, stressed the need for quantitative statements of stability of factors as age and group changes.

Psychologists have been wont to describe individuals on the basis of their test results. But they have not always been able to agree as to the specific qualities measured by the various tests. Murphy (19), in studying the relation between tests of mechanical ability and verbal and non-verbal tests of intelligence, found much confusion of terminology existing in the field of measurement of mechanical aptitude. Worthington (31) noted that the *Picture Completion*

---

[2] A fuller discussion of this scale will be given in Section II.

Test, Cancellation of A and *Knox Cube* were tests of special ability and should be further analyzed. Pintner and Anderson (20) described the *Picture Completion Test* as being a measure of the child's apperceptive ability, "to see how well he is able to meet the requirements of each situation by grasping its meaning." Louttit (16) has written:

> Just what abilities are measured in form-board performances is not at all clear. Various writers have claimed their usefulness in measuring form perception and discrimination, manipulative ability, motor coördination, etc., but there have been no attempts to establish adequately their ability in any of these fields.

The same difficulties in interpretation have been expressed with regard to other tests.

The research by means of factor analysis has made a great contribution toward the isolating of the mental factors involved in the tests of intelligence, and toward a description of the various tests in terms of these factors. Stephenson (23) found the verbal factor *v*, and Alexander (1) established as a separate factor in certain tests what he called *f*, a performance factor. Morris (18) recently analyzed "a group of 34 of the most commonly used performance tests" and found them to be describable in terms of only three factors, called visualization, perceptual speed, and induction. Wright (32) made a factor analysis of a part of the original Stanford-Binet Test, extending from the 7-year level to the 14-year level, and identified six factors, called number, space, imagery, verbal relations, induction, and one involving reasoning ability. Of course, Thurstone's analysis (27) of 56 psychological tests must be included in any account of factor analysis. As is generally known, he found seven factors, which he called primary abilities, among the 56 tests.

This study necessarily has its limitations. It does not attempt to isolate all the possible mental factors at each age level investigated, or to study each year period from nine to sixty. It will, however, extract and name as many factors as the test data will allow. Six representative age groups have been chosen for detailed analysis. These are the following: 9 years, 12 years, 15 years, 25 to 29 years, 35 to 44 years, and 50 to 59 years. An analysis of these age periods may serve to throw light on the problems under consideration.

II. DESCRIPTION OF SAMPLINGS AND TESTS

A. THE SAMPLINGS

As previously stated, the present study does not attempt to investigate each successive age year. Nor does it extend the investigation below the ninth year, since the scale was not designed to measure children below this age. At the ninth year significant scores, i.e., scores better than zero on any single test, could still be obtained. The age groups selected for study are, as indicated above: 9 years, 12 years, 15 years, 25 to 29 years, 35 to 44 years, and 50 to 59 years.

These various periods may be briefly characterized. Nine years is definitely a pre-pubescent age. Twelve years can be regarded as an age when puberty begins for many individuals. Fifteen years is that period where the curve of mental growth is known to approach a maximum. Asch (3) has studied the character of mental organization in the two age levels nine and twelve. In a study of the age factor in mental organization, Garrett, Bryan, and Perl (9) used the age levels 9, 12, and 15, which they stated were "far enough apart on the growth curve to make age a real factor." Twenty-five to 29 years is generally accepted as that period of life in which almost everyone has completed his education and training, and has embarked on some definite activity or career. This age period was considered the optimal age group in the standardization of the Wechsler-Bellevue Scale, and was used as a basis for comparison with other age periods. Thirty-five to 44 years can be considered as a period when there is the beginning of a noticeable decline in the curve of mental growth. The period from 50 to 59 years taps that part of the curve of mental growth and decline where the decline is definitely noticeable. This particular age group was selected in preference to any other of the older age groups, because more cases were available between those years.

All of the individuals used in the various samplings were selected from the files used in the standardization of the Bellevue Scale, and were made available by Dr. Wechsler. None of the subjects were patients in the hospital. All the subjects were tested by examiners who had become qualified to administer the scale after a training period under Wechsler. The investigator himself examined many of the individuals during the course of standardization of the scale.

Each age sampling was selected to comply as nearly as possible with Kelley's criteria of homogeneity from the standpoint of maturity, sex, race, and general scholastic training. This being a study of differential traits, it is necessary that as many factors as possible be held constant. The result must not be marred by variations which are attributable to extraneous factors. Homogeneity of samplings will avoid spurious correlations. Kelley has written on this point (14):

> The reader can readily satisfy himself that race heterogeneity, providing the individuals of the one race are on the average lower as adults in both traits considered than those of the other race, operates to introduce correlation just as do unequal levels of maturity. Similar observations apply to sex and nurture. We may, therefore, conclude that even if the selection of the two mental traits is excellent in the sense that the mental capacities are intrinsically independent, positive correlations will nevertheless commonly be found between them, owing to the heterogeneity of the population employed.

Although sex differences were statistically insignificant for any of the subtests of the Wechsler-Bellevue Scale used in this study, whatever differences might have still existed were removed by partialling out the factor of sex in every intercorrelation of two subtests for every age level. The various samplings were selected randomly within the requirements set forth by Kelley.

The group of nine-year-olds consisted of 70 cases ranging from nine years and zero months to nine years and eleven months. All the children were white, American-born, and attended one of three elementary schools located in middle-class sections of New York City.

The group of 12-year-olds consisted of 75 cases ranging from the age of 12 years and zero months to that of 12 years and 11 months. All the children were white, American-born, and attended one of three elementary schools located in middle-class sections of New York City.

The 15-year-olds comprised a group of 93 cases ranging from the age of 15 years and zero months to 15 years and 11 months. All the children were white, American-born, and attended one of three general high schools in middle-class sections of New York City.

The age group 25 to 29 years included 135 subjects. All were white, American-born, and had had some high school training, ranging from the first year to graduation. They were employed in occupations which were considered to be on a similar level in terms of the abstract intelligence required. They all belonged to Category IV of the "Six Categories of Abstract Intelligence" of the *Minnesota Occupational Rating Scale* described by Bingham (5). Some of the occupations described in this category are those of auto mechanic, stationary engineer, file clerk, and typist.

The age group 35 to 44 consisted of 121 individuals. All were white, American-born, and had had educational training ranging from the sixth year of elementary school to the first year of high school. This group was employed in occupations similar to those of the 25- to 29-year-old group. On the average, the older adult groups have not received as much education as the younger. Older individuals with less education have been employed in occupations similar to those of younger persons with more education.

The age group 50 to 59 years old consisted of 69 cases. All were white, and were either American-born, or else had been in the United States for at least 45 years. They had all attended school in the United States, and their education ranged from the fifth grade of elementary school to the eighth. They were also employed in occupations similar to those of the two younger adult groups.

No attempt was made to control racial differences in this study, because it was permissible to combine groups of differing racial origin. It was, of course, necessary to assume that these racial groups had, at any given age, the same mean level of attainment with regard to the traits in question. On this point, Wechsler has written (28): "We believe, however, that the differences that might be found between the English-speaking white groups in the United States of different 'racial' origin would not be very large." Kelley says (14): "Here in America we can throw together a number of racial stocks and approximate this condition" (equivalence of means at any age).

The groups selected had average *IQ*'s on the Wechsler-Bellevue Scale approximating 100 (Table A).

The agreement in average *IQ* is relatively good when compared with standardization groups where the average *IQ*'s range similarly

TABLE A

	Chronological age					
	9	12	15	25-29	35-44	50-59
Mean *IQ*	100.1	99.4	103.5	104.1	103.9	104.5
Standard deviation	14.19	13.44	11.42	11.18	11.32	13.45

from age to age.[3] On the basis of the closeness of average *IQ*'s to each other, as noted in the table above, it can be said that the groups used in this study are of similar intellectual level.

B. The Tests

The Wechsler-Bellevue Scale used in this study is an adolescent and adult test of intelligence which is individually administered. It is a point scale consisting of 10 different subtests, five verbal and five non-verbal, bearing the following names: (1) Information; (2) Comprehension; (3) Arithmetic; (4) Memory Span for Digits; (5) Similarities;[4] (6) Picture Arrangement; (7) Picture Completion; (8) Object Assembly; (9) Block Design; (10) Digit Symbol (Substitution).

The test of *Information* consists of 25 items arranged in order, from the least to the most difficult. Of this test Wechsler (28) has written: "The fact is, all objections allowed for, the range of a man's knowledge is generally a very good indication of his intellectual capacity." This test is one of the five verbal tests.

The test of *Comprehension* has 10 items, each scored on a two, one, and zero standard according to defined criteria. These items are also arranged in the order of their difficulty. About this test, Wechsler has written:

> Precisely what functions the Comprehension Test involves is difficult to say. Offhand it might be termed a test of common sense and it is so called on the Army Alpha. Success on the test seemingly depends on the possession of a certain amount of practical information and a general ability to evaluate past experience.

This is also considered a verbal test.

A test of *Arithmetical Reasoning*, consisting of 10 items ranged in order of difficulty, is the third test in the battery. This test was designed to be interesting to adults as well as to children, and is also included under the sub-heading of *Verbal Tests.* Of this particular test, Wechsler has written:

---

[3] L. M. Terman and M. A. Merrill, *Measuring Intelligence* (p. 35). In the table of *IQ* means by age for the standardization group, the means (from *Form L* of the late revision of the *Stanford-Binet Scale*) vary from a high of 109.9 at age two and a half to a low of 100.9 at the age of 14. In *Form M*, they vary from a high of 108.3 at the age of two and a half to a low of 101.2 at the age of 14.

[4] The similarities test is not included in this study, because it was introduced late in the standardization of the scale, and many of the subjects used in the various samplings had not been given this particular test (28, p. 9).

> The ability to solve arithmetical problems has long been recognized as a sign of mental alertness. Even before the introduction of psychometrics, it was used as a rough and ready measure of intelligence.

The fourth test of the verbal group is the *Memory Span for Digits.* The fifth, *Similarities,* was not included, as noted above. The *Memory Span for Digits* has both the forward and the backward series combined in one test, and so allows for a wider range of scores. In discussing this particular test, Wechsler has written: "The ability involved (in this test) contains little of '*g*' and as Spearman has shown is more or less independent of the general factor."

The first of the series of non-verbal tests is the *Picture Arrangement Test,* consisting of a series of pictures, which, when placed in the proper sequence, tell a story. There are six such sets of pictures that go to make up this subtest. The first four are scored for accuracy alone, and the last two for both accuracy and time. The last three of the picture series have more than one scorable solution. Wechsler noted concerning this test that:

> It is the type of test which effectively measures a subject's ability to comprehend and size up a total situation. The subject must understand the whole, must get the "idea" of the story before he is able to set himself effectively to the task.

The *Picture Completion Test* consists of a series of 15 pictures, each with some important part missing. The subject has to discover and name the missing part. This test, according to Wechsler, ostensibly "measures the individual's basic perceptual and conceptual abilities in so far as these are involved in the visual recognition and identification of familiar objects and forms."

The *Block Design Test* is a modification of the Kohs series, and consists of seven designs, none of which are found in that series. In the Wechsler-Bellevue Scale all the designs are in the two colors, red and white. Wechsler described this test as the "best single performance item," and it seems that the "reproduction of the type of design devised by Kohs in some way involves both synthetic and analytical ability."

The *Object Assembly Test* is not a single item. It contains three figure form-boards: a Manikin, a Feature Profile, and a Hand. The Manikin is the one made familiar by the Pintner-Paterson Scale. The Feature Profile is similar to that of the Pintner-Paterson Feature Profile, but is the profile of a female instead of a male, has a two-piece ear instead of a four-piece one, and has an additional cutout at the base of the skull. The Hand is a new figure form-board, which has been mutilated by cutting off the fingers and a large section of the palm. This test was included in the Scale because, according to Wechsler, "it tells us something about one's mode of perception, the degree to which one relies on trial and error methods, and the manner in which one reacts to mistakes."

The last test, the *Digit Symbol or Substitution Test,* is the familiar *Army Beta Digit Symbol Test.* The time allowed for this test has been reduced from two minutes to one and one-half. Wechsler has noted that this test requires the subject "to associate certain symbols with certain other symbols, and the speed and accuracy with which he does it serve as a measure of his intellectual ability."

III. PRESENTATION OF DATA

Each one of the age samplings has been treated separately. All possible correlations between subtests were calculated by Pearson's product-moment method, and the results were put into the form of correlation matrices. In order to eliminate completely any influence that might have been due to the grouping together of the sexes in the various samplings, each subtest was correlated against the sexes by the bi-serial method and the partial coefficients were calculated for the six correlation matrices.[5] These new sets of interest correlations were factored by Thurstone's centroid method, and are shown in Tables 1-6.

Analysis of the various correlational matrices was carried out to three factors for age nine, to four for ages 12, 15, 35 to 44, and 50 to 59, and to five factors for age 25 to 29. Of course one could continue to extract as many factors as there are test variables. But this would be an unscientific procedure, since the original purpose has been to reduce the number of variables. If one continued to extract factors up to the limit of the test variables, one would eventually be extracting artifacts due to sampling errors and the unreliability necessarily ascribable to any correlation figure.

The unrotated factors extracted from the correlation matrix

TABLE 1

AGE 9

Summary of intertest correlations (sex factor held constant)

	Com.	Inf.	Arith.	Digit span	P.A.	P.C.	Obj.	Bl. des	Subst.
Com.		.509	.310	.482	.392	.227	.412	.445	.220
Inf.	.509		.445	.460	.406	.305	.412	.416	.462
Arith.	.310	.445		.419	.205	.268	.244	.283	.276
Digit span	.482.	460	.419		.374	.048	.359	.421	.291
P.A.	.392	.406	.205	.374		.339	.471	.581	.387
P.C.	.227	.305	.268	.048	.339		.233	.238	.051
Obj.	.412	.412	.244	.359	.471	.233		.505	.322
Bl. des.	.445	.416	.283	.421	.581	.238	.505		.536
Subst.	.220	.462	.276	.291	.387	.051	.322	.536	

Probable errors range from .081 for $r = 0.00$ to .054 for $r = .581$.

---

[5] It is to be noted that partialling out the sex factor produced very slight changes in the original correlations, the maximum change being .08, thereby indicating that the sex differences were insignificant, as had already been pointed out by Wechsler.

TABLE 2

Age 12

Summary of intertest correlations (sex factor held constant)

	Com.	Inf.	Arith.	Digit span	P.A.	P.C.	Obj.	Bl. des	Subst.
Com.		.582	.476	.411	.230	.340	.292	.341	.458
Inf.	.582		.622	.485	.298	.283	.259	.399	.510
Arith.	.476	.622		.387	.130	—.026	.141	.267	.494
Digit span	.411	.485	.387		.293	.343	.004	.241	.495
P.A.	.230	.298	.130	.293		.321	.275	.259	.191
P.C.	.340	.283	—.026	.343	.321		.133	.306	.277
Obj.	.292	.259	.141	.004	.275	.133		.432	.264
Bl. des.	.341	.399	.267	.241	.259	.306	.432		.479
subst.	.458	.510	.494	.495	.191	.277	.264	.479	

Probable errors range from .077 for r = 0.00 to 0.48 for r = .622

TABLE 3

Age 15

Summary of intertest correlations (sex factor held constant)

	Com.	Inf.	Arith.	Digit span	P.A.	P.C.	Obj.	Bl. des	Subst.
Com.		.395	.345	.153	.245	.288	.140	.282	.248
Inf.	.395		.419	.131	.204	.276	.189	.358	.108
Arith.	.345	.419		.157	.449	.149	—.019	.247	.165
Digit span	.153	.131	.157		.088	—.019	.116	—.018	.014
P.A.	.245	.204	.449	.088		.352	.190	.168	.173
P.C.	.288	.276	.149	—.019	.352		.345	.353	.252
Obj.	.140	.189	—.019	.116	.190	.345		.415	.100
Bl. des.	.282	.358	.247	—.018	.168	.353	.415		.082
Subst.	.248	.108	.165	.014	.173	.252	.100	.082	

Probable errors range from .070 for r = 0.00 to .057 for r = .449.

TABLE 4

Age 25-29

Summary of intertest correlations (sex factor held constant)

	Com.	Inf.	Arith.	Digit span	P.A.	P.C.	Obj.	Bl. des	Subst.
Com.		.421	.208	.030	.186	.103	.075	.239	.188
Inf.	.421		.255	.033	.284	.236	.135	.229	.191
Arith.	.206	.255		.226	.190	.188	—.104	.070	—.011
Digit span	.030	.033	.226		.003	.062	—.112	—.017	.227
P.A.	.186	.284	.190	.003		.180	.059	.363	.134
P.C.	.103	.236	.188	.062	.180		.324	.305	.025
Obj.	.075	.135	—.104	—.112	.059	.324		.514	.168
Bl. des.	.239	.229	.070	—.017	.363	.305	.514		.205
Subst.	.188	.191	—.011	.227	.134	.025	.168	.205	

Probable errors range from .058 for r = 0.00 to .043 for r = .514.

TABLE 5

AGE 35-44

Summary of intertest correlations (sex factor held constant)

	Com.	Inf.	Arith.	Digit span	P.A.	P.C.	Obj.	Bl. des	Subst.
Com.		.535	.308	.290	.152	.329	.173	.267	.281
Inf.	.535		.471	.374	.245	.259	.282	.260	.364
Arith.	.308	.471		.344	.281	.397	.320	.246	.411
Digit span	.290	.374	.344		.169	.153	.160	.145	.362
P.A.	.152	.245	.281	.169		.408	.289	.425	.334
P.C.	.329	.259	.397	.153	.408		.322	.332	.185
Obj.	.173	.282	.320	.160	.289	.322		.538	.162
Bl. des.	.267	.260	.246	.145	.425	.332	.538		.357
Subst.	.281	.364	.411	.362	.334	.185	.162	.357	

Probable errors range from .061 for $r = 0.00$ to .042 for $r = .538$.

TABLE 6

AGE 50-59

Summary of intertest correlations (sex factor held constant)

	Com.	Inf.	Arith.	Digit span	P.A.	P.C.	Obj.	Bl. des	Subst.
Com.		.629	.352	.169	.445	.383	.238	.406	.451
Inf.	.629		.547	.307	.472	.411	.171	.519	.355
Arith.	.352	.547		.399	.566	.390	.387	.620	.479
Digit span	.169	.307	.399		.353	.306	.182	.297	.306
P.A.	.445	.472	.566	.353		.587	.466	.565	.553
P.C.	.383	.411	.390	.306	.587		.451	.635	.427
Obj.	.238	.171	.387	.182	.466	.451		.466	.558
Bl. des.	.406	.519	.620	.297	.565	.635	.466		.548
Subst.	.451	.355	.479	.306	.553	.427	.558	.548	

Probable errors range from .081 for $r = 0.00$ to .049 for $r = .635$.

TABLE 7

AGE 9

Centroid factor loadings and communalities as computed by Thurstone's centroid method

Subtest	I	II	III	Communality h^2
1	.640	—.214	.285	.537
2	.717	—.223	—.198	.603
3	.529	—.354	—.220	.454
4	.609	—.172	.282	.480
5	.682	.253	.107	.541
6	.374	—.102	—.137	.169
7	.632	.140	.134	.437
8	.732	.325	.070	.646
9	.563	.276	—.234	.448
ΣK^2	3.432	0.547	0.356	
$\frac{\Sigma K^2}{N}$	.381	.061	.039	

TABLE 8

AGE 12

Centroid factor loadings and communalities as computed by Thurstone's centroid method

Subtest	I	II	III	IV	Communality h^2
1	.704	.116	.055	.103	.523
2	.770	.245	.035	.158	.679
3	.591	.506	—.248	.191	.703
4	.599	.259	.295	—.242	.572
5	.440	—.274	.204	.108	.322
6	.440	—.268	.388	—.200	.456
7	.424	—.375	—.296	.212	.453
8	.608	—.281	—.269	—.174	.551
9	.698	.137	—.181	—.304	.631
ΣK^2	3.221	0.782	0.538	0.350	
$\frac{\Sigma K^2}{N}$	.358	.087	.060	.043	

TABLE 9

AGE 15

Centroid factor loadings and communalities as computed by Thurstone's centroid method

Subtest	I	II	III	IV	Communality h^2
1	.581	—.134	—.072	—.091	.369
2	.583	—.150	—.226	—.267	.485
3	.557	—.492	.133	—.146	.585
4	.182	—.221	—.214	.283	.208
5	.541	—.104	.361	.126	.450
6	.548	.311	.181	—.041	.432
7	.441	.439	—.226	.233	.493
8	.537	.274	—.189	—.295	.486
9	.324	.052	.215	.005	.154
ΣK^2	2.195	0.709	0.416	0.340	
$\frac{\Sigma K^2}{N}$	.244	.079	.046	.038	

are presented in Tables 7-12. On studying these factors for each age level, it is found that the third factor extracted for age nine has no factor as high as .316 (the communality of which is its square and therefore about .100). Age 12 has no factor loading in the fourth factor with a communality as high as .100. The third factor has one loading as high as .388 and several approaching .316. The same

TABLE 10

Age 25-29

Centroid factor loadings and communalities as computed by Thurstone's centroid method

Subtest	I	II	III	IV	V	Communality h^2
1	.484	.290	.173	.180	.114	.394
2	.572	.323	.121	.062	.167	.478
3	.330	.312	—.376	—.056	—.078	.357
4	.176	—.023	—.362	.326	—.226	.320
5	.457	.153	.091	—.172	—.299	.360
6	.453	—.126	—.192	—.269	.190	.367
7	.408	—.495	.252	—.208	.240	.576
8	.628	—.297	.252	—.278	—.172	.653
9	.351	—.161	.113	.421	—.152	.362
ΣK^2	1.780	0.685	0.502	0.545	0.334	
$\frac{\Sigma K^2}{N}$	.200	.076	.056	.061	.037	

TABLE 11

Age 35-44

Centroid factor loadings and communalities as computed by Thurstone's *centroid method*

Subtest	I	II	III	IV	Communality h^2
1	.562	.246	—.278	.242	.512
2	.651	.348	—.130	.201	.602
3	.636	.171	.059	—.181	.470
4	.464	.318	.139	.071	.341
5	.534	—.285	—.073	—.266	.443
6	.547	—.203	—.238	—.170	.426
7	.545	—.360	.201	.248	.529
8	.608	—.392	.158	.175	.579
9	.561	.191	.210	—.189	.431
ΣK^2	2.925	0.754	0.288	0.364	
$\frac{\Sigma K^2}{N}$	.325	.084	.032	.040	

observations can be made for age 15. The fourth factor of age 25 to 29 has loadings over .316, but the fifth has none at all. Age 35 to 44 presents a more difficult problem. Here both third and fourth factors have no loadings of the order .316. But since the age group 50 to 59 has three factors that are significant, it was felt that the three-dimensional system should be retained for the age group 35

TABLE 12

AGE 50-59

Centroid factor loadings and communalities as computed by Thurstone's centroid method

Subtest	I	II	III	IV	Communality h^2
1	.617	—.536	—.066	—.078	.679
2	.673	—.340	.336	.026	.682
3	.726	.100	.191	.280	.652
4	.453	.096	.165	.259	.309
5	.765	.093	—.068	.053	.601
6	.704	.204	.011	—.338	.652
7	.579	.217	—.370	—.063	.523
8	.781	.136	.121	—.185	.677
9	.705	—.001	—.339	.098	.622
ΣK^2	4.089	0.538	0.452	0.317	
$\frac{\Sigma K^2}{N}$	.454	.060	.050	.035	

TABLE 13

AGE 9

Psychologically meaningful factor loadings obtained after rotation

Subtest	*A*	*B*	Communality h^2
1	.491	.460	.453
2	.560	.499	.563
3	.332	.541	.403
4	.481	.412	.401
5	.722	.060	.525
6	.293	.247	.147
7	.620	.139	.404
8	.801	.009	.642
9	.626	—.017	.392
ΣK^2	2.921	1.007	
$\frac{\Sigma K^2}{N}$	.325	.112	

to 44. Further justification for this practice was offered by a consideration of the fact that, if the third factor were really significant, it would show up as such in the rotations.

The unrotated factor loadings were then rotated for each age period to make them psychologically meaningful. This was done by the graphical method described by Guilford (10). Two factors are dealt with at one time, "rotating them about a third axis as a

TABLE 14

AGE 12

Psychologically meaningful factor loadings obtained after rotation

Subtest	A	B	C	Communality h^2
1	.587	.346	.209	.508
2	.728	.298	.196	.657
3	.778	—.123	.220	.669
4	.613	.365	—.084	.516
5	.128	.517	.156	.308
6	.128	.633	.015	.419
7	.038	.253	.586	.409
8	.244	.320	.602	.524
9	.601	.184	.382	.541
ΣK^2	2.311	1.325	1.016	
$\frac{\Sigma K^2}{N}$	.257	.136	.113	

TABLE 15

AGE 15

Psychologically meaningful factor loadings obtained after rotation

Subtest	A	B	C	Communality h^2
1	.245	.473	.273	.358
2	.142	.488	.395	.414
3	.166	.729	—.038	.560
4	—.126	.285	.168	.125
5	.497	.424	—.076	.433
6	.603	.117	.229	.430
7	.357	—.054	.557	.441
8	.357	.140	.507	.404
9	.356	.168	—.001	.155
ΣK^2	1.116	1.319	.886	
$\frac{\Sigma K^2}{N}$	.124	.145	.098	

pivot, computing new factor loadings after each rotation, and repeating the process as many times as is necessary"—that is, until further rotations no longer decrease the number of negative loadings. The rotated loadings are presented in Tables 13-18. All of the factors in each age group are psychologically meaningful and, being orthogonal to each other, are each independent of the other.

The subtests of the Wechsler-Bellevue Scale are listed by number

TABLE 16

AGE 25-29

Psychologically meaningful factor loadings obtained after rotation

Subtest	A	B	C	D	Communality h^2
1	.540	.059	.270	.071	.373
2	.504	.018	.415	.125	.442
3	.034	.182	.551	—.127	.354
4	.051	.496	.126	—.106	.276
5	.244	—.087	.381	.246	.273
6	—.085	.115	.403	.379	.326
7	.025	.001	—.033	.729	.539
8	.174	—.048	.231	.743	.639
9	.395	.364	—.101	.208	.342
ΣK^2	.809	.439	.937	1.379	
$\frac{\Sigma K^2}{N}$	.089	.049	.104	.153	

TABLE 17

AGE 35-44

Psychologically meaningful factor loadings obtained after rotation

Subtest	A	B	C	Communality h^2
1	.197	.219	.605	.453
2	.397	.213	.601	.564
3	.441	.333	.368	.441
4	.483	.099	.299	.333
5	.068	.580	.171	.370
6	.001	.530	.347	.401
7	.241	.643	—.038	.473
8	.226	.707	.004	.551
9	.519	.262	.235	.393
ΣK^2	1.007	1.813	1.158	
$\frac{\Sigma K^2}{N}$	.112	.201	.129	

in the various Tables. They are: (1) Comprehension; (2) Information; (3)Arithmetical Reasoning; (4) Digit Span (Forward and Backward); (5) Picture Arrangement; (6) Picture Completion; (7) Object Assembly; (8) Block Design; (9) Digit Symbol (or Substitution). The communalities are presented to show how much of the total variance of each test is accounted for by the various factor loadings. The Communality is merely the sum of the squares of the factor loadings for each test, regarded as percentages. The com-

TABLE 18

AGE 50-59

Psychologically meaningful factor loadings obtained after rotation

Subtest	A	B	C	Communality h^2
1	.068	.804	.171	.680
2	.474	.675	.001	.680
3	.582	.359	.328	.575
4	.414	.191	.184	.242
5	.415	.391	.531	.607
6	.479	.261	.482	.530
7	.151	.173	.689	.527
8	.579	.357	.428	.646
9	.151	.427	.646	.622
ΣK^2	1.522	1.828	1.760	
$\frac{\Sigma K^2}{N}$	.169	.203	.196	

munality minus 100 per cent represents the uniqueness of the test, i.e., the specific factor and chance and other factors due to possible errors of unreliability. The sum of the squares of the factor loadings of each test is represented by ΣK^2 and $\frac{\Sigma K^2}{N}$ indicates the relative percentages of each factor, that is, the amount of variance which is attributable, on the average, to the various factors.

IV. INTERPRETATION OF THE RESULTS

It is evident, upon investigation of the intertest correlational matrices, that the correlations between the subtests of whatever age group are positive, except for a few which, although negative, are insignificantly so. The number of negative correlations are fewer in the original correlation matrix than in the matrix with the sex factor eliminated. But they are in each case sufficiently close to zero to be called insignificantly negative. That most of the correlations are positive is to be expected, since each of the several subtests was included in the Scale because it was purported to measure intelligence or some attribute of it. The rather low order of the various intercorrelations is in line with the facts found in all mental tests. A test of intelligence has low correlations between its parts, but each part correlates to a much higher degree with the entire scale. Wechsler (28) gives the correlations of each part of his Scale with the total. They are all much higher than the intercorrelations among the parts.

But it is also evident that there are changes in the degree of correlation from age to age. The 36 intercorrelations calculated at each age level were averaged in order to note these changes better (Table B).

TABLE B

	Chronological age					
	9	12	15	25-29	35-44	50-59
Average of the "*r*"s	.37	.34	.23	.18	.31	.43
Standard deviation	*.106*	*.128*	*.111*	*.110*	*.100*	*.126*

There is, then, a definite decrease in the average of the subtest intercorrelations through the age group 25 to 29, followed by a decided increase for the older ages. This suggests a greater differentiation in abilities with probable specialization, and a later reorganization of abilities due perhaps to complete maturation of all functions, which would seem to indicate a greater flexibility in the handling of complexities.

The tendency for intercorrelations of tests to decrease with age has been noted by Garrett, Bryan, and Perl (9), who wrote concerning their study "that the rôle of general ability is minimized in favor of special abilities as age increases." Previous investigations on other groups of children by Thorndike (24), by Bryan (7), and by Anastasi

(2), have led to essentially similar findings. Asch (3), in a study of change in mental organization, stated that:

> The results of the present investigation show that between the ages of 9 and 12 there has occurred in our group a significant reduction in the relationships between a number of intellectual functions. The factor analysis confirms this finding and further indicates that a considerable portion of the reduction has occurred in the factor of "general ability."

Most of the literature on this subject deals with the child of school age, but some of it is concerned also with the preschool child. Furfey, Bonham, and Sargent (8) administered a scale of 17 items to 462 new-born infants, ranging in age from 15 minutes to 343 hours. They concluded that "there is no mental integration in the newborn." Bayley's (4) later work studied infants from birth to three years of age, each child receiving at least six tests. On the basis of her findings, Bayley concluded:

> That the behavior growth of the early months of infant development has little predictive relation to the later development of intelligence—even though the later behavior may depend in large part on the previously matured elementary neural connections or behavior patterns.

Very few such analyses have been made of the adult. Jones and *Conrad (13) have described some significant changes in performance* on the Army Alpha of subjects ranging in age from 10 to 60 years. They reported that:

> In the sixth decade of life, about 40 per cent of total Alpha score is derived from two tests (Tests 4 and 8); at age 10, these tests contribute only 25 per cent. As represented in mental tests, then, the effective intellectual power of the adult, much more than that of the 10-year-old, is evidently derived from accumulated stocks of information.

Another investigation, this one dealing with physical traits, offers data strikingly in agreement with those found in the present study. Reed and Love (21) examined changes in interrelationships of a large number of physical measures in a group of 5,000 United States Army officers. These officers had been examined yearly and such measures as stature, weight, chest circumference, pulse rate, and systolic and diastolic blood pressures had been taken. Changes were found in the interrelationships between physical traits. The correlations between diastolic and systolic blood pressure should be noted (Table C).

TABLE C

	Chronological age				
	21-25	28-35	36-45	46-55	56-64
Correlation	.26	.48	.61	.73	.74
Probable error	.039	.013	.009	.008	.014

It becomes evident, then, that the same abilities are not tapped by tests of intelligence at various age levels, or that the abilities themselves are not constant. It remains to be seen what light the analysis of the mental factors can throw on this problem.

At this point it would be well to name the isolated, independent factors for each of the several age groups. First to be considered is the age group nine. For purposes of determining which of the rotated loadings of the various subtests are significant in each factor, use has been made of the criterion that the individual loading had to "exceed three times the standard error of an original correlation of zero," in the words of Woodrow (30). For age nine the factor loading must exceed .359.

Reference to Table 13 will give the rotated loadings for age nine. Significant loadings for the first factor are:

Comprehension = .491
Information = .560
Digit Span = .481
Picture Arrangement = .722
Object Assembly = .620
Block Design = .801
Digit Symbol = .626

It would seem that this factor corresponds to Spearman's *G*, since it has significant loadings in seven of the tests and rather significant loadings in the remaining tests of *Arithmetic* (.332) and *Picture Arrangement* (.293).

Significant loadings for the second factor are:

Comprehension = .450
Information = .499
Arithmetic = .541
Digit Span = .412

This seems to be the factor of verbality or verbalization. It might also be the factor of memory, since each test involves that type of function. But for age nine it can be said that the ability demonstrated in these tests is not so thoroughly routinized as to require merely memory.

For age 12, reference should be made to Table 14. The significant loadings at this age must exceed .347. For the first factor they are:

Comprehension = .587
Information = .728
Arithmetic = .778
Digit Span = .613
Digit Symbol = .601

This can be said to be a verbal factor. The inclusion of the Digit Symbol Test in this factor is probably explainable by the fact that this type of task is common in the school system, and develops along with other verbal functions.

The second factor in age 12 has the following significant test loadings:

Comprehension = .347
Digit Span = .365
Picture Arrangement = .517
Picture Completion = .633

This factor very probably is concerned with the seeing of relationships in social situations. That *Digit Span* is also included in this factor would indicate that memory for social situations is required before one can see relationships in them.

Factor three in age 12 has the following significant test loadings:

Object Assembly = .586
Block Design = .602
Digit Symbol = .382

This factor can be identified broadly as a performance factor. The very high loadings in the *Object Assembly* and *Block Design Tests* may lead one to believe that the factor is that of visualization, involving perception of form and space relations. That the *Digit Symbol Test* contains such a factor is quite possible.

Table 15 gives the rotated factor loadings for age 15. At this age a significant factor loading exceeds .311. The first factor has significant loadings in the following tests:

Picture Arrangement = .497
Picture Completion = .603
Object Assembly = .357
Block Design = .357
Digit Symbol = .356

This may best be identified as a non-verbal or performance factor. The second factor has significant loadings in the following tests:

Comprehension = .473
Information = .488
Arithmetical Reasoning = .729
Picture Arrangement = .424

This may be best described as a verbal or verbalization factor. Digit Span has a rather high loading here too, being equal to .285. That the *Picture Arrangement Test* falls in this group is not surprising, since verbalizing the sequence is quite common.

The third factor has the following significant test loadings:

Information	= .395	Block Design	= .507
Object Assembly	= .557		

What the test of *Information* has in common with the other two tests is not very clear, except perhaps that they all involve some form of restriction accompanying their solution. But this is not likely to be so, since the test of *Arithmetical Reasoning* has a negative loading in this factor. It is very difficult to describe this factor, except to say that the tests all require an awareness to the environment.

In ages 25 to 29 a significant loading exceeds .258. By reference to Table 16 it will be seen that the first factor has significant loadings in the following tests:

Comprehension	= .540	Digit Symbol	= .395
Information	= .504		

This is best described as a verbal or verbalization factor. That the *Digit Symbol Test* is grouped along with the other two at this age level indicates that this particular type of work is quite clerical in nature and has developed along with verbal ability. It is to be noted that, at this age, the tests of *Arithmetical Reasoning* and *Digit Span* have almost zero loadings in this factor, giving evidence of specialization in these abilities.

The second factor has significant loadings in these tests:

Digit Span	= .496	Digit Symbol	= .364

This appears to be a memory factor, since each test certainly involves that function.

The third factor has the following significant test loadings:

Comprehension	= .270	Picture Arrangement	= .381
Information	= .415	Picture Completion	= .403
Arithmetical Reasoning	= .551		

This particular grouping seems best explainable in terms of a restriction in solution. All the tests seem to involve that factor.

The fact that *Arithmetical Reasoning* and *Information* involve this factor, more than do the other tests, affords reasonably good evidence that it is the factor called restriction by Thurstone.

The fourth factor of the age group 25 to 29 contains significant loadings in these tests:

Picture Completion = .379   Block Design = .743
Object Assembly = .729

This can be described as a performance factor, or one that deals with spatial material. The tests of *Picture Completion* and *Digit Symbol,* with the almost significant factor loadings of .246 and .208 respectively, tend to bear this out. All the remaining tests, considered to be verbal in nature, have either approximately zero or low negative loadings in this factor.

In the age group 35 to 44 a significant loading exceeds .273. On referring to Table 17 it will be seen that the first factor contains significant loadings in the following tests:

Information = .397   Digit Span = .483
Arithmetical Reasoning = .441   Digit Symbol = .519

This factor evidently calls upon memory, since all tests are known to involve it. That *Information* has a significant loading in memory at this age is quite understandable. That *Arithmetical Reasoning* at this age involves so much memory is quite noteworthy.

The second factor has significant loadings in the following tests:

Arithmetical Reasoning = .332   Object Assembly = .643
Picture Arrangement = .580   Block Design = .707
Picture Completion = .530

Digit Symbol with a loading of .262, is rather significant. It is evident that all five tests considered to be of a performance nature are, then, included in this one factor. Why *Arithmetical Reasoning* is also part of this group is not clear, except that this test has a significant loading in all three of the factors. It is rather likely that all tests involve reasoning of some kind. It is known that the so-called performance tests become more difficult for the older age groups, as shown by an investigation of the norms for the various tests supplied by Wechsler in the text on his Scale.

The third factor has significant loadings in these tests:

Comprehension = .605   Digit Span = .299
Information = .601   Picture Completion = .347
Arithmetical Reasoning = .368

This factor is probably best described as dealing with verbal relations. The two highest loadings are in the tests of *Comprehension* and *Information,* and this fact definitely points to such an interpretation. It is not clear why *Picture Completion* also has a significant loading in this factor, but it may be noted that this test and that of *Picture Arrangement* seem to slip easily into the verbal groups.

In the age group 50 to 59 a significant loading exceeds .361. Reference to Table 18 shows that the first factor has significant loadings in the following tests:

Information = .474
Arithmetical Reasoning = .582
Digit Span = .414
Picture Arrangement = .415
Picture Completion = .479
Block Design = .579

This is a peculiar admixture of verbal and non-verbal tests, indicating a breakdown in this type of categorization for the older age group. All that one can say about such a grouping of tests is that they all contain some form of reasoning. It is to be noted that *Digit Span* is included in this factor, and at this age may require the application of some sort of reasoning.

The second factor has significant loadings in the following tests:

Comprehension = .804
Information = .675
Picture Arrangement = .391
Digit Symbol = .427

Arithmetical Reasoning, with a factor loading of .359, and *Block Design,* with a loading of .357, are also fairly significant. Here again it is extremely difficult to interpret the grouping in terms of a single factor. All the tests have positive loadings, and the average variance of all the tests is the highest of the three factors extracted for this group, so that this factor may be similar to the *G* factor found in the nine-year group.

The third factor of the age group 50 to 59 years has the following significant test loadings:

Picture Arrangement = .531
Picture Completion = .482
Object Assembly = .689
Block Design = .428
Digit Symbol = .646

This is apparently a performance factor. It is to be noted that the *Arithmetical Reasoning Test* has a fairly significant loading in this group, as also in the age group 35 to 44 years.

On looking over Tables 13-18 of *Rotated Factor Loadings,* it may be observed that a single subtest may have significant loadings in one or more factors. This indicates that the subtests are complex rather than simple. This complexity is characteristic of most tests that go to make up an intelligence scale, and is observed in the factor analyses of other tests and scales. This makes the interpretation of these tests rather difficult, and helps to account for the apparently opposing, or at least different, interpretations of the exact qualities measured by any specific test.

Although certain factors, such as the verbal, performance, restriction, and some sort of reasoning factors, occur as independent and may be isolated several times throughout the various ages, the same subtests do not always compose them. It has been noted that the tests of *Picture Arrangement, Picture Completion,* and *Digit Symbol* fall into groups that are known to be verbal tests. There are times when not all the so-called performance tests fall together to give significant loadings in the performance factor. The factors themselves change in the average percentage of variance attributable to them from age to age. For instance, the verbal factor at age nine has an average variance of 11 per cent; at age 12 it is 26 per cent; at age 15, 15 per cent; at age 25 to 29, 9 per cent; at age 35 to 44, 12 per cent; and at age 50 to 59 there is no clearly defined verbal factor.

It is to be noted that the test of *Digit Span* is in the group called verbal at ages *9, 12,* and *35* to *44*, but is not in the verbal group at age 15, and comes out as a separate factor at age 25 to 29. The same phenomenon may be noted with regard to the test of *Arithmetical Reasoning*. The performance factor does not appear until age 12, where it does not contain the tests of *Picture Arrangement* and *Picture Completion*. At age 15 this factor contains all performance tests; at age 25 to 29 it contains all performance tests, but with two having only fairly significant loadings. At age 35 to 44 the performance factor contains the test of *Arithmetical Reasoning* besides the performance tests, and at age 50 to 59 the performance factor also has a fairly significant *Arithmetical Reasoning Test* loading. All these data tend to indicate that the factors, although recurrent, take on different meaning in terms of the various subtests as they go from age to age. Any generalized statement about test results must, then, take into account the factor of age.

A careful study of both the unrotated factor loadings and the rotated, psychologically meaningful factors also reveal much pertinent material. It will be necessary to examine the first factor of unrotated loadings for each age level. This first factor necessarily accounts for most of the variance of the subtests, because the factor method is such that the first extraction takes out most of the variance of the tests. For age nine it accounts for 38 per cent of the total variance; for age 12, 36 per cent; for age 15, 24 per cent; for ages 25 to 29, only 20 per cent; for ages 35 to 44, 33 per cent; and for ages 50 to 59, 45 per cent. The magnitude of these first unrotated loadings is unaffected by further extractions, and may thus be compared inter se. It is immediately evident, then, that less and less of the variance can be accounted for by a single factor through the age group 25 to 29, while more and more of the variance can be so accounted for as the higher age groups are reached. This finding fits in perfectly well with the fall and rise of the average intercorrelations as noted above, and lends strength to the interpretation that there is a greater specialization up to a certain point, followed by a later reintegration of the various abilities into a flexible whole.

The older adult, above the age of about 27½ years, becomes more able to utilize his various abilities for different tasks. It is as though the various specializations became part of the individual, and he was now able to use them readily and flexibly throughout more than one task. As the immature person makes use of several abilities in order to solve a particular problem for which he as yet has no special method of organization, so does the older, completely mature adult. But the young person was forced to use several abilities because he had no specialized method for attacking the problem, while the older person made use of a variety of abilities, because all his abilities were integrated into such an organization that he could use them easily and flexibly.

The general intelligence factor *G*, described by Spearman as pervading all mental tasks, can be said to appear, after rotation, in the age groups 9 and 50 to 59. Inspection of Tables 13-18 containing the rotated loadings, shows that the highest single average percentage of the factor loadings varies from a high of 33 per cent at age nine, to a low of 15 per cent at ages 15 and 25 to 29, and up again to 20 per cent at 50 to 59. Although the age groups 12 and 35 to 44 also have high average percentages, 26 per cent and 20 per cent respectively, the individual loadings for the various subtests indicate descriptions on less general terms. Thus the ages 12 and 35 to 39 have more definitely insignificant subtest loadings in the factor accounting for the highest average percentage of variance than do the ages nine and 50 to 59.

On looking at the figures headed $\frac{\Sigma K^2}{N}$ in the Tables 13-18 of Rotated Loadings, one can once more note the trend toward greater specalization up through the age 25 to 29, and then, for the later ages, less specialization indicating greater complexity. The figures $\frac{\Sigma K^2}{N}$ show the amount of variance, actually described in percentages, which is attributable, on the average, to the various factors. Smaller percentages are noted up to the age 25 to 29, and then larger ones appear again. Thus for age nine the percentages are, respectively, 33 per cent and 11 per cent for factors one and two; and for age 12, 26 per cent, 14 per cent, and 11 per cent for factors one, two, and three. At age 15, the percentages are lower, being 12 per cent, 15 per cent, and 10 per cent respectively for factors one, two, and three. They are still lower for age 25 to 29; here the percentages are respectively 9 per cent, 5 per cent, 10 per cent, and 15 per cent for factors one, two, and three, and four. At age 35 to 44 they rise once more to 11 per cent, 20 per cent, and 13 per cent for factors one, two, and three, and at age 50 to 59 they are as high as 17 per cent, 20 per cent, and 20 per cent respectively for the three factors.

V. SUMMARY AND CONCLUSIONS

In order to study the mental factors and to note any changes in these factors and their organization, various age samplings spreading over the years from nine to sixty were selected. The respective age samplings are 9 years, 12 years, 15 years, 25 to 29 years, 35 to 44 years, and 50 to 59 years. All of these age groups had been given the same test, namely the Wechsler-Bellevue *Intelligence Scale.*

The subtests that compose the Scale were each correlated with the other at every age level, the sex factor completely removed from the intercorrelations by partialling it out, and the new intercorrelations of the subtests factored by Thurstone's centroid method. These factor loadings were then rotated until they gave psychologically meaningful factors or abilities, each independent of the other.

The naming of the factors, of course, required thorough study of the tests that had significant loadings in them. The factors could only be named after much speculation, and at times could not be definitely described. It is, nevertheless, important to note that the names given to the factors by all the investigators employing the factor analysis technique are similar, and all the factors found in this study have been named by other research workers, using different samplings and different tests.[9]

The independent mental factors found in this study are the following:

1. For age nine, a *G* factor and a verbal factor.
2. For age 12, a verbal factor, a performance factor, and one called seeing relationships in social situations.
3. For age 15, a verbal factor, a performance factor, and one that could not be clearly indicated.
4. For age 25 to 29, a verbal factor, a performance factor, a memory factor, and a factor called restriction in solution.
5. For age 35 to 44, a verbal factor, a performance factor, and a memory factor.
6. For age 50 to 59 a *G* factor, a performance factor, and a factor involving some sort of reasoning.

It should be emphasized at this time that the factors isolated in this study are dependent on the subtests used. If one were to use different tests, one would naturally expect different factors. Again, if other tests were added to the battery of tests used in this scale, other groupings might become evident. It can be said, however, that the above factors are the ones existent in the Wechsler-Bellevue Scale, and are sufficient to describe all the test variables at each level.

---

[9] Thus Thurstone has identified such factors as Verbal, Memory, Induction, and Restriction, among others. Alexander has identified such factors as *G*, verbal, and performance.

But this study was undertaken with the intention of doing more than merely discovering and isolating the mental factors. The purpose of the investigator was to note the stability of these factors from age to age, to study the tests that entered into the factors at each age level, and to obtain further facts regarding mental organization.

Concerning the stability of the factors from age to age, the following conclusions have been reached:

1. That the same factors do not always appear at each age level.
2. That the verbal and performance factors are most consistent.
3. That the memory factor appears only in ages 25 to 29 and 35 to 44.
4. That a G factor is found for the nine-year age group, then is apparently submerged and appears again, but not as definitely, at age 50 to 59.

On looking over that part of Section IV where the factors were named, it will be seen that the same tests did not always enter into the same factor at each age level. It can also be seen that where the same tests did enter into the same factors, they did not always do so to a similar degree.

The study has yielded significant facts about mental organization. This matter has been more thoroughly discussed in the preceding section. The average of the subtest intercorrelations ranged from a high of .37 for age 9 to a low of .18 at age 25 to 29, and then back to a high of .43 for age 50 to 59. The first factor loadings extracted at each age level also indicated a similar trend. Thus, for age 9, the first factor accounts for 38 per cent of the total variance of the subtests; for age 12, 36 per cent; for age 15, 24 per cent; for age 25 to 29, only 20 per cent; for age 35 to 44, as much as 33 per cent; and for age 50 to 59, almost half, or 45 per cent. The magnitude of these first factor loadings is unaffected by further extractions, and these are comparable among themselves.

Summarizing the trends indicated in the data, it can be said that there was noted a tendency toward greater specialization from age 9 to age 25 to 29, and thereafter an apparent reorganization to a complexity which could be described as flexible.

General conclusions gleaned from all the data, including the matrices of intercorrelations, the unrotated factor loadings, and the rotated factor loadings, are as follows:

1. The subtest intercorrelations were all positive, except for a few at ages 12, 15, and 25 to 29, which were very low negative.
2. The average of the subtest intercorrelations decreased from age 9 to 25 to 29, and then increased up to age 50 to 59.
3. The percentage of variance due to the first of the centroid factor loadings showed a similar trend from an original high of 38 per cent at age 9 to a low of 20 per cent at age 25 to 29, and then back to another still greater high of 45 per cent at age 50 to 59.
4. Certain descriptive factors could be isolated at the various ages, but did not always have the same tests grouped within them.
5. Individual subtests changed their factorial composition from age to age.
6. The average variances due to the various rotated factor loadings were greater for ages 9, 12, 35 to 44, and 50 to 59 than they were for ages 15 and 25 to 29.
7. A *G* factor seemed to be present at ages 9 and 50 to 59, but was not disclosed for the other age groups.
8. It was evident from an inspection of the communalities of the various subtests that each subtest had a unique character in addition to the group characters, since none of the communalities approached unity.
9. As a result of the above findings, it could be stated that the mental traits change and undergo reorganization over a span of years. Therefore, when interpreting tests of intelligence, it is of importance to take into consideration the age of the individual. The same test, given to a person of a certain age, may not be measuring the same abilities in him that it would measure when given to an older or younger person. Even though the whole intelligence scale may yield the same factors for a wide span of years, the separate tests that compose the scale may not necessarily be described in terms of the same factors from age to age.

REFERENCES

1. Alexander, W. P. Intelligence, concrete and abstract. *Brit. J. Psychol. Monog. Suppl.*, 1935, No. 19. Pp. 177.
2. Anastasi, A. A group factor in immediate memory. *Arch. of Psychol.* 1930, No. 120. Pp. 61.
3. Asch, S. E. A study of change in mental organization. *Arch. of Psychol.*, 1936, No. 195. Pp. 30.
4. Bayley, N. Mental growth during the first three years: A development study of sixty-one children by repeated tests. *Genet. Psychol. Monog.*, 1933, **14**, 1-92.
5. Bingham, W. V. D. Aptitudes and Aptitude Testing. New York: Harper, 1937.
6. Brown, W., & Thomson, G. F. The Essentials of Mental Measurement. London: Cambridge Univ. Press, 1925.
7. Bryan, A. Organization of memory in young children. *Arch. of Psychol.*, 1934, No. 162. Pp. 56.
8. Furfey, P. H., Bonham, M. A., & Sargent, M. K. The mental organization of the newborn. *Child Devel.*, 1930, **1**, 48-51.
9. Garrett, H. E., Bryan, H. J., & Perl, R. E. The age factor in mental organization. *Arch. of Psychol.*, 1935, No. 176. Pp. 31.
10. Guilford, J. P. Psychometric Methods. New York: McGraw-Hill, 1936.
11. Heidbreder, E. Seven Psychologies. New York: Appleton-Century, 1933.
12. Hotelling, H. Analysis of a complex of statistical variables into principal components. *J. Educ. Psychol.*, 1933, **24**, 417-441, 498-520.
13. Jones, H. E., & Conrad, H. S. Growth and decline of intelligence: A study of a homogeneous group between the ages 10 and 60. *Genet. Psychol. Monog.*, 1933, **13**, 223-298.
14. Kelley, T. L. Crossroads in the Mind of Man. Stanford University: Stanford Univ. Press, 1928.
15. ———. Mental factors of no importance. *J. Educ. Psychol.*, 1939, **30**, 139-142.
16. Louttit, C. M. Clinical Psychology. New York: Harper, 1936.
17. McCloy, C. H., Metheny, E., & Knott, J. A comparison of the Thurstone method of multiple factors with the Hotelling method of principal components. *Psychometrika*, 1938, **3**, 361-367.
18. Morris, C. M. A critical analysis of certain performance tests. *J. Genet. Psychol.*, 1939, **54**, 85-105.
19. Murphy, L. W. The relation between mechanical ability tests and verbal and non-verbal intelligence tests. *J. of Psychol.*, 1936, **2**, 353-366.
20. Pintner, R., & Anderson, M. M. The Picture Completion Test. Baltimore: Warwick & York, 1917.
21. Reed, L. J., & Love, A. G. Biometric studies on U. S. army officers: Somatological norms, correlations, and changes with age. *Human Biol.*, 1932, **4**, 509-524.

22. SPEARMAN, C. The proof and measurement of association between two things. *Amer. J. Psychol.*, 1904, **15**, 72-101.

23. STEPHENSON, W. Tetrad difference for verbal subtests relative to non-verbal subtests. *J. Educ. Psychol.*, 1931, **22**, 334-350.

24. THORNDIKE, E. L. Measuring human intelligence. *Harper's Magazine*, 1930.

25. THURSTONE, L. L. Multiple factor analysis. *Psychol. Rev.*, 1931, **38**, 406-426.

26. ———. Vectors of Mind. Chicago: Univ. Chicago Press, 1935.

27. ———. Primary mental abilities. *Psychomet. Monog.*, 1938, No. 1.

28. WECHSLER, D. Measurement of Adult Intelligence. Baltimore: Williams & Wilkins, 1939.

29. WILSON, E. B., & WORCESTER, J. Note on factor analysis. *Psychometrika*, 1939, **4**, 133-148.

30. WOODROW, H. The common factors in fifty-two mental tests. *Psychometrika*, 1939, **4**, 99-108.

31. WORTHINGTON, M. R. A study of some commonly used performance tests. *J. Appl. Psychol.*, 1926, **10**, 216-227.

32. WRIGHT, R. E. A factor analysis of the original Stanford-Binet scale. *Psychometrika*, 1939, **4**, 209-220.

Published as a separate and in *Genetic Psychology Monographs*, 1941, 23, 237-318.

FACTORS INFLUENCING PERFORMANCE ON GROUP AND INDIVIDUAL TESTS OF INTELLIGENCE: I. RATE OF WORK*

MARY WOODS BENNETT

Institute of Child Welfare, University of California

- Acknowledgments 239
- I. Introduction 241
  - *A.* General discussion 241
  - *B.* Factors influencing performance on group and individual tests 244
- II. Review of related studies 251
  - *A.* Theoretical considerations 251
  - *B.* The relation of speed to intelligence 256
- III. The experiment 271
  - *A.* The problem 271
  - *B.* The sample 272
  - *C.* Instruments and apparatus 275
  - *D.* Procedure 277
- IV. Results 281
  - *A.* The data 281
  - *B.* Speed and altitude in intelligence test performance . 295
- V. Summary and conclusions 303
- Appendices 305
  - Appendix *A* 305
  - Appendix *B* 310
- References 311

---

*Accepted for publication by Harold E. Jones of the Editorial Board, and received in the Editorial Office on December 1, 1940. Copyright by The Journal Press.

Assistance in the preparation of these materials was furnished by the personnel of Works Progress Administration Official Project No. 65-3-5406.

ACKNOWLEDGMENTS

The writer is deeply indebted to those whose generous assistance made this investigation possible: to Dr. Harold E. Jones, Director of the Institute of Child Welfare, for making available the facilities of the Institute and for guidance and constructive criticism in every phase of the study; to Dr. Herbert S. Conrad, for helpful suggestions as to the statistical analysis of the data and the organization of the material; and to the Oakland Public Schools, for coöperation in the testing program. Grateful acknowledgment is further made of the many services rendered by members of the Institute staff.

MARY WOODS BENNETT

Institute of Child Welfare
University of California
Berkeley, California

I. INTRODUCTION

The literature on mental testing contains many expressions of opinion on the part of experts as to the relative merits of group and individual examinations. Analysis of the various influences which might affect performance on the different types of test reveals several factors concerning which additional information is needed.

One of the common objections to the use of the ordinary group test of intelligence is that it penalizes the slow but able thinker by placing too high a premium upon mere speed of performance. This point of view has led to investigations of the nature of speed and its relation to general intelligence, but results are conflicting and further study is required. The present investigation is an attempt to review the work which has been done in this field and to present the results of an experiment designed to throw some light on the questions raised by earlier workers.

A. General Discussion

Any discussion of intelligence tests and testing—whether it be in article, manual, or textbook—usually presents to the reader a series of contrasts between types of examinations. Verbal tests[1] are shown to differ from non-verbal, those of the pencil-and-paper variety from the performance type, point scales from mental age scales. Occasionally the distinction is made between speed and power tests. By far the most common contrast is the one between testing procedure in individual as compared with group examinations.

Individual tests, as the name implies, are those which are administered by the examiner to one person at a time, usually without an audience, whereas group tests are those which are given to a number of subjects in the same space-time. Authors who discuss the matter usually concede that each type of instrument has its place, but the clinical psychologist is likely to stress the greater thoroughness of individual examinations, while the educator is inclined to emphasize the saving of time and expense by the use of group tests. This difference in emphasis on the part of two large classes of test-users is readily explained by the divergent needs and purposes of the two groups.

---

[1] For descriptions of these and other types of tests see References 17, 20, 47, 53, and 103.

The clinician employs tests of general intelligence[2] as part of his analysis of individual problems of adjustment. His aim is to collect as much and as accurate information as possible about a person's capacities, special proficiencies and deficiencies, physical and emotional development, educational status, and the like. He uses his findings as the basis for individual diagnosis and recommendation. The planning of programs for the adjustment of difficulties, and decisions as to the desirability of institutional placement, are often in his hands. Hence the intensive training necessary to equip the examiner and the time involved in testing persons one at a time do not constitute for the clinician legitimate objections to the individual type of mental test. More important to him are the opportunities offered in this situation to control testing conditions, to supplement the quantitative measure of intelligence with a qualitative analysis of the responses made, and to obtain additional information about the person tested through observation during the examination period.

The educator, on the other hand, uses intelligence tests for group comparisons, for rough classification of students, or as a means of discovering those individuals who should be examined more intensively. Under these circumstances, and because of the large numbers usually involved, group tests are to be preferred. Their application to whole classes at once, by persons who need not be highly trained, secures economy of time and money.

Authorities in the field of mental measurement (28, 34, 82, 103, 104, 112) are agreed in the main on two points: namely, that more different sorts of information are gained by individual tests and that such examinations are for the individual case more trustworthy. That the first of these two propositions is true can hardly be denied. The nature of the responses made, aside from their correctness, and information over and above that yielded by either the quantitative or qualitative results of the test itself are valuable data usually obtainable only through individual examination. In the earlier days of intelligence testing both Bronner (16) and Kent (76) warned against the use of the bare quantitative measure in estimating mental ability, and the latter author pointed out that the individual test situation offers an opportunity for recording of incidental notes and for observation of the following aspects of behavior: coöperation and interest, attention, persistence, susceptibility to fatigue, emotional manifestations, defects of speech and coördination, nervousness, adaptability to test conditions, comprehension of instructions, rate of work, and methods of work. In more recent years Wires (166), Zachry and Lloyd (171), and Kent (78) have sounded the same warning. Recognition of the significance of these supplementary data has led psychologists to advocate the use of individual tests in industrial (156) and professional (73) selection; consideration of the qualitative aspects of performance has proved fruitful in studies of psychotics (31, 76, 115), defectives (97, 113), and delinquents (64). Furthermore it has been pointed out that individual examinations permit the use of more tests and of a wider variety (15, 87, 88, 158).

---

[2] This expression is here used in the common-sense meaning of the term and implies no conviction on the part of the writer as to the nature of "general intelligence" or as to the theoretical adequacy of the tests designed to measure it.

We turn now to a discussion of the second proposition upon which most authorities agree, that is, that individual tests are somehow better than group tests. Individual examinations are described by various authors as more valuable (122), more dependable (68), more trustworthy (114), more accurate (34, 41, 90, 104, 140, 170), more satisfactory (100), and a purer measure of intelligence (112) than the group type. Group tests are said to present greater chances of erroneous results (152). The difference between the two kinds of tests is more precisely stated by those who claim greater validity (88, 103) and reliability (33, 60, 67, 103, 135, 136) for the individual examination. Although the majority opinion favors the individual tests, there have been some dissenting voices. Burt (22) was among the first of these. He suggested that group tests given at different schools on the same day at the same time by teachers of sciences who had a knowledge of experimental psychology might yield results more comparable with one another than those obtained by a group of experts testing children individually at varying times and seasons under varying conditions of fatigue and familiarity. Franzen (44) in 1922 concluded from a study of test validation that the average of several group tests might be an adequate substitute for the Binet test as a measure of intelligence. At about the same time Gates (57) and Geyer (58) were investigating the reliability of group tests and reporting it as inadequate; more recently Nemzek (101) and Miller (98), in the course of studies on the constancy of the *IQ*, have reached the conclusion that reliability of group tests as measured by correlation between repeated examinations compares favorably with that of the individual type. McCall (90), while recognizing the advantages of individual tests, points out that some subjects may be hindered rather than helped by the more personal relation with the examiner which obtains under conditions of individual testing. Furthermore he states that a group test may yield fairer results for a school population by removing the opportunity for coaching or of additional learning which would benefit those tested later in an individual testing program. Wells (162) suggests that group tests, being the result of coöperation among many experts and capable of wider application, are generally better organized and more carefully standardized, that their scores depend less on the skill of the examiner, hence that they may often be a better means than the individual tests of obtaining an intelligence score. Few are so bold as to say with Symonds (138, p. 187), "there is evidence to support the belief that group tests are fully as satisfactory as individual tests."

The fact that the conclusions concerning the relative merits of group and individual examinations are often based on comparisons between tests of different content, administered under varying conditions to non-comparable samples, serves to confuse a problem that is already complex. Lack of agreement among authors as to definitions of reliability and validity, their use of the terms interchangeably and with several different meanings, further complicate the issue. In order to simplify future discussion, the "reliability" of a test will be taken to mean the consistency with which it measures a trait or capacity (52, p. 268), and "validity" will be used to designate the degree to which a test measures what it is supposed to measure (53, p. 44). It remains now to clarify certain aspects of the comparison between group and individual tests by analysis of factors which may influence performance on the two types of examination, thus affecting their relative reliability and validity.

B. Factors Influencing Performance on Group and Individual Tests

1. *Control of Testing Conditions*

Among the factors cited by authorities to demonstrate the superiority of individual examinations, those which may be classed together under the general heading of control of testing conditions are most frequently mentioned. It is pointed out that in testing one person alone the examiner may more often secure optimum conditions of external environment, such as appropriate lighting and ventilation, a room free from distraction, comfortable chair and desk for the subject, and so on. The effect of chance occurrences such as the breaking of a pencil need not be reflected in the test score. Furthermore, the physical condition of the subject, major sensory or motor defects, evidences of malnutrition or fatigue, minor disorders such as headache or cold, and general state of health may all be noted and taken into account. The examination may be abandoned or postponed, or interpreted in the light of these findings. One of the important advantages of the individual testing situation is the possibility of establishing "rapport," that relationship between examiner and examined which sets the subject at his ease and insures his best effort in the tasks at hand. Finally, the effect on the test performance of chronic or acute emotional disturbances may often be detected and given weight in the appraisal of the final results.

Not only are these opportunities lacking, or present to a slighter extent, in group testing, but the possibility of copying from one's neighbors, of failing to observe time limits, and of distraction by the presence of others constitute further disadvantages. Lest the case for the group test be made to seem hopeless, it is well to remember Pintner's (112, p. 180) contention that the errors in group testing due to the sources here described have been overemphasized.

Presumably this better control of testing conditions makes possible a more reliable measure of intelligence because it permits the examiner to eliminate, or at least to be aware of and to take into account, some of the factors which might lead to variations in score of the same person on the same test at different times. That more valid measures of intelligence are yielded by individual tests seems likely in view of the fact that the examiner is able to prevent or ameliorate the influence of certain conditions which, while not affecting the reliability of the examination, still might prevent the subject from giving a performance indicative of his ability. An apprehensive child, for example, might be rated consistently low on group tests but might, under the more favorable conditions of rapport in individual testing, show a consistently superior score.

In this discussion an effort is made to distinguish between the advantages derived from control of conditions and those resulting from the supplementary observation and qualitative analysis of individual performance. Both procedures are ultimately concerned with the validity of the final estimate of the subject's intelligence. Each represents an effort to obtain as pure and as complete a measure as possible, but the former is in general concerned with prevention of vitiation of the results by irrelevant influences and the latter with enriching the measure. In so far as these procedures are separable, and it is admitted that the boundary between them is not clean-cut, the concern of the present study is with the preventive procedure of control of conditions.

It is granted that enrichment of the measure is desirable and probably attainable to the fullest extent only through individual testing. It is now pertinent to inquire whether the control of conditions as a preventive device is necessarily exclusive to individual examinations. Practically the answer must be in the affirmative, for by definition the group test involves one examiner and many subjects, but theoretically it is conceivable that either by reducing the size of the group or by multiplying examiners, group testing could be made to approach the conditions of individual testing. How far this process might be carried, and how close the approximation to individual conditions might be, need not concern us here. The significant point is that, given competent examiners, securing optimum conditions seems to depend on the number of examiners in relation to the number of persons examined and not on the individual situation as against the group situation. The adequate control of test conditions may be more difficult under group conditions but it is not theoretically impossible. The importance of this distinction will be revealed later.

2. *Age of Person Tested*

The discussion just concluded may be extended by mention of a further factor influencing performance on group and individual tests of itelligence, that is, the age of the person examined. Control of conditions in group testing is more difficult with young children. Very young subjects cannot be depended upon to follow instructions given to the group as a whole, nor are there many tests suitable for little children which can be administered to more than one at a time. Hence results of group tests applied at the lower age levels are apt to be unreliable and invalid. On the other hand, it has been suggested (77) that older subjects, adolescents and adults, find the individual testing situation embarrassing, and it is possible therefore that they do their best and most consistent work on group examinations. Comparisons between the two types of test might favor one or the other according to the age level of the persons examined.

3. *Test Content*

Up to this point there has been no clear distinction made between individual tests and the individual testing situation. The great majority of references to individual examinations in the literature concern those of the Binet type, usually the Stanford Revision (139, 143), or the so-called performance tests of intelligence. References to individual testing usually imply testing with material of the Binet or performance type, whereas discussions of group testing ordinarily assume the use of group examinations of the pencil-and-paper variety. It has been shown time and again (32, 38, 55) that performance on most of the standard group tests is more influenced by reading skill than is performance on the *Stanford Revision*. It is known that the non-verbal skills and abilities measured by performance tests or non-verbal group examinations are not always closely related to the verbal ability which is a large factor in success on the Stanford Revision or on the verbal group tests (7, 55, 56, 66, 171). It is conceivable that reliability coefficients of two tests differing in content, one administered individually and one administered to a group, might be influenced both by the nature of the material and by the method of administration. Furthermore, the relative validity of two such tests would depend, at least in part, on the opinion of the examiner as to the nature of intelligence or on the character of the external criterion of intelligence set up.

As was suggested earlier, typical individual tests differ from typical group tests not only in the kind of material but also in the variety and often in the amount of material. Hence part of the superior reliability claimed for individual examinations may result from longer tests, part of the superior validity from a more comprehensive sampling of those functions which go to make up intelligent behavior. It is apparent that unless the factor of test content is controlled, differences in the results of individual and group testing procedures may be subject to varying interpretations.

4. *Social Facilitation*

Whipple (164) suggested the existence of another factor influencing results on group and individual tests of mental ability when he pointed out that there was no way of telling how the necessity of working with others might affect the performance of a given subject. Others have intimated that some persons might be stimulated by working in a group but have not indicated in detail the nature of the influences presumably operative under such conditions (13, 124, 159). It was Allport (4) who gave the name of social facilitation to the factor leading to increased efficiency of group work. He described it as "an increase of response merely from the sight or sound of others making the same movements" (4, p. 262), and distinguished it from rivalry, "an emotional reinforcement of movement accompanied by the consciousness of desire to win" (4, p. 262). Recently Dashiell (30, p. 1111) has stated the problem as follows:

> An eminently practical question arises in connection with the administration of intelligence tests. These have been constructed in the two forms of individual and group tests, and the former have been considered the more valid, but on account of the better opportunity to fit the test items to the individual examinee and other methodological and mechanical advantages and not on account of any belief that the examinee works better alone. Now, it is pertinent to ask, does he work on an intelligence test better in the one situation than in the other?

It is conceivable that if a factor of social facilitation does exist, group test scores might be higher than individual test scores for the same person, and possibly a better indication of his true ability. Such a factor, present in repeated group test situations but lacking in individual examinations, might lead to more consistent performance, hence to more reliable results, on group examinations.

In a discussion of the facilitating influence of the group on performance it is necessary also to consider the possibility of impeding effects. The disruptive influence of the group on test performance has been suggested by many (71, 129, 159), and Allport (3) himself points out that the group situation may foster such impeding factors as distraction, over-rivalry, and emotional disturbance. It is his conclusion, however, that such effects are relatively less important than those of social facilitation. Mention was made, in the discussion of control of conditions, of the possibility of distraction caused by the presence of others in the group testing situation. It seems likely that such disturbances are not necessarily dependent upon the group procedure itself but that they are akin to the noises, interruptions, and other nuisances which may beset the individual examination as well. To the extent that control of group conditions may be made to approach that of individual conditions, such effects should be capable of reduction. The impeding factors which remain might be referred to as the negative aspect of social facilitation and should be considered in any attempt to investigate the latter.

5. *Rate of Performance*

Whipple (164) was among the first to introduce the question of rate of performance in a discussion of individual and group examinations. It was his conclusion that speed was not an adequate indicator of general mental efficiency, hence that the use of group tests, which had to be administered with time limits, was to be deprecated. Later writers either adopted the "slow but sure" hypothesis (33), namely that slow but able thinkers were penalized by the time limits necessitated by the group procedure, or brought forward evidence to show that the slow mind was also the inferior mind (41, 129), hence not unduly hampered by such restrictions. The advantages of the "unhurried response" in the individual situation have been stressed by some (46, 79, 112), the possible stimulating effect of the speed conditions of group testing by others (129).

Clearly the significant question here is not that of the superiority of individual over group procedures, for although group tests are usually administered with time limits, and individual examinations, at least of the Binet type, usually lay less stress upon rate of work, still there is nothing to prevent group testing with ample time allowance nor individual testing with strict time limits. The fundamental issue is that of the relationship between speed of response and intelligence. If it be true that there is a close relation between the two, then there can be little objection to group examinations on the score of their emphasis on rate of work. If, on the other hand, any negative relationships are found, the validity of intelligence tests which put a premium on speed could be challenged. That this problem has been recognized is abundantly proved by the welter of material published on the subject. That conclusions as to the nature of the relationship have been reached is evidenced by the construction of such intelligence examinations as the *CAVD*[3] and by the distinctions made between tests of this type and the "speed" tests with relatively easy items and strict time limits (23, 34, 79, 103). That the problem is not exhausted nor the conclusions final is demonstrated by the conflicting results reported by investigators and by the continued use of those tests which are said to emphasize rate of performance.

This discussion of influences which may lead to different results on examinations under group and individual conditions of administration has brought out certain points which help to clarify the question which prompted it—that of the relative merits of individual and group tests of intelligence. The problem seems to be not so much one of choosing between the two types of test as of determining what influences are operative with a given test under given conditions. To summarize briefly, it is suggested that the superior reliability and validity claimed for a particular type of test may be due not only to the method of administration but also to a combination of other influences such as adequacy of control of testing conditions, the content of the examination, and the age of the persons tested. The only factor which appears to depend directly and fundamentally upon the group or individual situation is that of social facilitation. The problem of whether tests necessitating time limits are inferior to those which may be given without time limits awaits further investigation of the relation between speed of performance and intelligence. The last-named factor has been selected for investigation, but before the problem is stated in final form a review of previous studies on the point will be presented.

---

[3] This is a group test consisting of items of graded difficulty and designed for administration without time limits. For further description see (144).

II. REVIEW OF RELATED STUDIES

A. Theoretical Considerations

It will be well to mention briefly at the outset the various trends in theory underlying the studies reviewed in this section. The experimental work which has been carried on has either been based, deliberately or unintentionally, on one or another of several assumptions, or it has been definitely designed to support or refute the different theories advanced. Consequently a better understanding of the implications for this study of the findings of previous investigators will follow from a consideration of the theoretical background of their work.

One of the important broad distinctions made in research on rate of performance is that between speed as a general personality trait and speed as a specific factor depending on the ability of the individual to meet the demands of a given situation. In the more limited field of the relation of speed to intelligence, the contrast of interest to us is that made between the doctrine which considers speed of mental performance as one with and inseparable from mental ability, and the one which conceives of speed as separate from, but possibly related to, intelligence.

1. *Speed as a Personality Trait*

The widespread popular belief in a general speed trait characterizing the behavior of a given individual has found support in the work of several investigators. The assumptions underlying the construction and use of the Downey *Will-Temperament Test* are typical. Downey (36) selected speed of reaction as one of the aspects of personality which should be taken into consideration in attempts to measure certain innate dynamic patterns of disposition. She believed that various aspects of the speed of response could be detected in samples of handwriting obtained under prescribed conditions, and she saw evidence of the existence of such a general character trait in the data derived from application of her test. More recently Kennedy (75) has claimed, on the basis of a review of earlier work and of her own findings on college students, that a personality trait exists which may be called "irritability," that it is not dependent upon intelligence, and that this trait, together with skill, determines rate of work in a specific task. Moore (99) analyzed the performance of 25 high school students on a variety of measures of linguistic, non-linguistic, and mechanical ability and concluded that there was indication of a function responsible for speed on all the tests.

Frischeisen-Köhler (51) postulated the existence of a "personal tempo" on the basis of responses to a series of tapping tests and metronome rates presented to subjects of various ages. Extending her study to include twin pairs and parent-child groups, this author interpreted her findings as indicating that personal tempo is innate, and she set up hypotheses to explain the biological mechanism by which it is inherited. Foley (42, 43), on the basis of an extensive review of the literature and an application of tapping and metronome tests to over 600 young women in a vocational high school, has challenged Frischeisen-Köhler's conclusions. Finding significant differences in motor speed and preferential auditory tempo for the various occupational groups and obtaining negative results as regards race differences in speed, he holds that vocational stimulation and similar environmental factors determine in large measure both maximal and preferential speeds. Allport (5), in a general discussion of personality traits, points out that the factor of psychic tempo must be in part innate if it exists at all, but he concludes that few persons appear to show a consistent tempo in all actions.

Evidence to support the theory of the specific nature of speed of response is more plentiful. In the wake of Downey's studies have followed investigations such as those of Meier (93), Trow (151), and Uhrbrock (155) which have demonstrated that whatever is measured by the speed items of the Will-Temperament tests is not clearly recognized as a trait by raters, and that any one of the types of speed—for example, speed of decision—is not always consistent for the same person in different tasks. As early as 1902, Aikens, Thorndike, and Hubbell (2) were presenting evidence to show that there was no such thing as a trait of quickness of association characterizing the work of a given individual on various simple mental tasks. Twenty-five years later investigations along the same line were being carried out by Dowd (35) and Sisk (128). The former worked with speed as measured in such performances as cancellation, handwriting, and hand movement; the latter used measures similar to these and also some derived from simple reaction time and more complex processes, such as looking up names in a directory. Both came to the conclusion that there is little evidence of a general speed factor, and Dowd suggested that high correlations among some of the speed measures were likely to be due to a similarity of material rather than to a characteristic rate of performance on the part of the person tested.

Among the more recent expressions of authoritative opinion on the subject are the conclusions of Allport and Vernon (6) drawn from a study of individual differences in expressive movement in 25 men. On the basis of ratings on speed and 13 different measures of speed of movement, these authors state that there is no conclusive evidence in support of a theory of general rate of work as a characteristic of personality. They found, instead, a tendency for verbal, drawing, and rhythmic speeds to correlate more highly among themselves than with each other. Subsequent studies in expressive movement, such as that of Carlson (24), who tested 100 college men on tasks similar to those used by Allport and Vernon, tend in general to verify the results obtained by the earlier investigators. It will be noted that these studies on the specificity of rate of work are concerned with all types of performances from simple reaction time to fairly complex mental tasks but that the emphasis has been on speed of movement and the simple intellectual functions. It is also true, as Allport and Vernon point out, that intercorrelations of about the same magnitude among speed measures have been used as evidence both for and against a general speed factor.

Before drawing to a close this brief treatment of the concept of rate of performance as a general personality trait and of the evidence in support or refutation of the theory, it will be interesting to take note of some of the work done in the field of race differences. Klineberg (81) found that white children were faster than Negroes and Indians on various performance tests and concluded that differences in speed attitude built up in the varying cultures to which the children belonged could be responsible for these results. The notion of a general speed factor dependent upon acquired rather than upon inherited traits and serving to differentiate members of cultural groups follows from this line of reasoning. Peterson (108), and Peterson, Lanier, and Walker (110) had earlier advanced a similar explanation for the relatively poor performance of Negro children on group tests of intelligence, and had found that speed scores on individual tests of reasoning also favored whites in the main. Further work by Peterson and Lanier (109) led to confirmation of the hypothesis. It remained for Lambeth and Lanier (83), working with 12-year-olds, to point out that when Negroes and whites were tested in intelligence, rational learning, and simple speed functions, the degree of difference in performance between the two races was so variable that it was not possible to assume a general difference in speed.

2. *Intellectual Speed as Identical with or Distinct from Intelligence*

Spearman (132) has been the chief proponent of the notion that speed of response in a mental task is a true indication of intelligence. Starting with the assumption that performance in intellectual work depends upon a combination of two kinds of factors—general intellective ability and specific factors peculiar to given tasks[4]—he has been faced with the necessity of demonstrating that there are no *additional factors common to all or several tasks and distinguishable* from general ability or *g*. This has been done in the case of speed by postulating clearness and speed of response as the quantitative properties of cognition knowable through goodness (accuracy) of the response and the amount of time taken to make it.

It is Spearman's contention (132, p. 246 ff.) that goodness and speed of response are interchangeable measures of *g*. To demonstrate the relationship between accuracy and *g* he cites high correlations between "measures of *g*" and scores on tests where speed of response is *not a factor in success. Furthermore, the existence of a* correlation of .66 between fineness of pitch discrimination and an accuracy score in addition is attributed to *g* and taken as evidence for the dependence of goodness of response on general intellective ability.

By correlational analysis of scores on such speed tests as addition and cancellation, Spearman finds evidence for the presence of *g*, and he concludes that this factor enters into speed of response. Correlations between speed scores in one sort of test and accuracy scores in another *are presented, and it is argued that since accuracy is* correlated with *g*, speed must be as well. Finally, the work of Courtis on the inverse relation found under certain conditions between speed and accuracy in a given task is interpreted as proof that both depend on the same factor, hence that if either speed or goodness of response depends on *g* the other one does also. The general conclusion is that speed and goodness are equal in respect to their dependence on general intellective ability, but it is admitted that in practice one of the two is often emphasized more than the other.

---

[4] Since no other explanation of the two-factor theory is possible here the reader is referred to the following sources listed in the bibliography (131, 132, 133, 134).

In his more recent comments on the subject Spearman (134) has distinguished between speed ability and speed preference. He claims that whereas a general speed preference may be revealed in performance of mental tasks, no general speed ability has been demonstrated. However, this holds only for eductive processes. Spearman is willing to admit that in such tasks as rate of tapping or rate of reacting there may be one or several speed factors independent of *g*.

The work of Spearman has stimulated many to follow his lead in attempting to demonstrate the existence of traits through mathematical analysis of correlational data. The fact that some of these support the two-factor theory and some deny it is not so significant for us as is the fact that, in denying it, some investigators have pointed to the existence of a speed trait or group factor. Kelley (74) interpreted his analysis of correlations as showing the existence of several independent traits characterizing mental tasks, speed in mental processes being among them. Line and Kaplan (85), and DuBois (37) have brought forth evidence to show the existence of a speed factor other than *g*. Thurstone (146) postulates seven primary mental abilities one of which, perceptual facility, seems to be related to speed of response.

Thorndike (144) is perhaps the best known and most outspoken supporter of the theory of specific traits. Not only has he been at pains to show, as was mentioned above, that speed is not a general trait possessed in greater or less degree by a given individual, but he also has submitted the proposition that speed, range, and altitude are separable aspects of mental ability. Altitude he defines as the level of difficulty which the individual can attain;[5] range, as the number of tasks a person can encompass at any particular level of difficulty; speed, as the individual's rate of performance. Granting that an intellect which can accomplish the largest number of tasks at the highest level of difficulty in the shortest period of time is in a very real sense the best intellect, Thorndike points out that altitude is the only indispensable aspect of intelligence and that it is imperfectly, though positively, correlated with speed. Thurstone (145) has more recently presented a theoretical discussion of the relation of ability, motivation, and speed in which he conceives of ability as independent of the other two factors.

---

[5] Henceforth when the terms intelligence or general mental ability are used they will be understood to refer to altitude in Thorndike's meaning of the word.

The physiological basis for the various factors and traits described in the foregoing paragraphs has not always been considered by those who approach the matter from the psychological or statistical point of view. Tryon (153) has pointed out that those who think of altitude and speed as separate and distinct abilities usually assume different physiological substrata for the two, such as number of connections for altitude and rate of nerve conduction for speed. On the other hand, those who assume partial or complete identity of the two imply their dependence upon the same physiological counterpart. Tryon has shown that this could be either rate of nerve conduction or number of connections.

We shall return to these theoretical assumptions in shaping the problem for investigation after consideration of work directly bearing on the issue.

B. The Relation of Speed to Intelligence

The review of the literature deals only with those studies which have as their avowed purpose the investigation of this relationship. There have been many which have yielded incidental information on the subject but which have been organized for the collection of other data, usually of wider application. These include investigations of the physical and mental traits of various populations (25, 59, 92, 167, 168), researches on motor and mechanical ability and on their relation to mental ability (8, 9, 10, 14, 105, 107, 116, 117, 125, 157), and numerous studies concerned with the two-factor theory of Spearman (1, 19, 62, 63, 137, 161, 169) or the more general aspects of mental testing (21, 22, 40, 72, 80, 95, 102, 118, 127). Many of the authors referred to in the discussion of speed as a personality trait have commented on the relation between speed and intelligence but with conclusions on the whole no different from those to be reported here. Investigations in the more restricted field of the relation of speed to intelligence will be treated under the following sub-headings: (1) reflex time and intelligence, (2) reaction time and intelligence, (3) speed on simple motor and mental functions as related to intelligence, (4) speed on simple intelligence test material as related to altitude, (5) speed and intelligence as measured by the same instrument.

1. *Reflex Time and Intelligence*

By means of the electromyograph, Travis and Hunter (149) were able to measure the time elapsing between stimulation of the patellar tendon and the appearance in the extensor muscles of the thigh of the electrical disturbance which sets off the muscular contractions of the knee jerk. This measure of speed was obtained on a group of 44 adults including college students, store and bank clerks, stenographers, housemaids, and farm hands, and on another sample composed of 43 unselected first year college men. Scores on the Otis *Higher Examination, Form A,* were available for the first group; on the University of Iowa qualifying examinations for the second group. Correlations between intelligence test scores and speed of reflex conduction, measured in thousandths of a second, were .87±.02 for both samples. Reliability of the intelligence test was given as .92 for the Otis, .95 for a composite of four parts of the Iowa. No coefficient of reliability was given for the speed scores, but the authors pointed out that a rereading of 15 of the records by a disinterested person showed almost perfect agreement. Travis and Hunter came to the conclusion that these results showed the importance to intelligence of speed of conduction of the nerve impulse and that they lent support to Spearman's notion of a physiological counterpart for speed of performance and *g*.

Travis and Young (150) later found evidence that this earlier work had been in error. Using a similar technique, but extending its application to 131 children at the preschool, second, third, fifth and sixth grade levels as well as to 122 university students, they found correlations between patellar reflex time and scores on various intelligence tests to be insignificantly different from zero.[6] Results were the same in comparisons between time on the Achilles reflex and measures of intelligence for 119 university students. The *Iowa Entrance Examination, Iowa Comprehension Test* and Stanford and Kuhlmann revisions of the Binet tests were used to measure intelligence in the various groups, and reliabilities of the first two were reported as .95 and .85 respectively. Reliabilities of the measures of the Achilles and patellar reflexes in college students were reported as .93 and .86 for a subgroup of six, while the reliabilities of the readings of the records were .97 and .94. The authors concluded that there was no evidence of a relationship between reflex time and intelligence, hence no reason for assuming a correlation between mental ability and speed of conduction of the nerve impulse in the reflex arc.

---

[6] In this and subsequent discussions a correlation or a difference is considered significant if it is at least four times its probable error. See (52, p. 136 and p. 170).

Travis and Dorsey (148) confirmed these findings in a study of reflex conduction rate in 57 inmates of an institution for the feebleminded. They found no significant differences in reflex time between these subjects and children of the same ages measured by Travis and Young, and they concluded that the findings of Travis and Hunter had been due either to the influence of uncontrolled factors or to the unreliability of the method. Whitehorn, Lundholm, and Gardner (165), using a technique somewhat similar to Travis', found no evidence for correlation between slow patellar reflex time and low mental age in a comparison between 13 feeble-minded and 13 normal adults. Neither in this experiment nor in those of Travis and his associates were the findings materially altered by taking account of the relationship between reflex time and height.

It is to be noted that the studies just mentioned have all been concerned with the relationship between speed of reflex response and altitude as determined either by measurement on a standard test of mental ability or by the classification of subjects in broad groups according to the degree of their intelligence. Rounds (119) concluded that there was a general speed ability rather closely associated with level of mental ability, but he based his arguments on six raw correlations ranging in magnitude from .21 to .51 between latent time in the Achilles reflex and scores on speed tests such as addition, completion, easy opposites, and cancellation. His reasoning appears to be that since these latter measures are indicative of speed in a learned response they are therefore symptomatic of ability to learn, or level of intelligence. It seems at least as likely, however, that what he has obtained are only correlations between different speed measures, and his assumption of a general speed factor cannot be seriously entertained in the absence of intercorrelations among the six measures of so-called mental speed. Rounds gives no coefficient of reliability for these measures which were collected on 80 college men.

In summary it can be stated that there is no conclusive evidence of a relationship between reflex time and mental ability as indicated by score on standard tests of intelligence or by coarser distinctions between feebleminded and normal subjects. As Tinker (147) has pointed out, however, there is as yet no accurate knowledge of how speed of conduction of the nerve impulse in the mechanisms operating in the reflex compares with that in the central nervous system. Hence there is no evidence of the relation between intelligence and the rate of conduction in complex mental processes.

2. *Reaction Time and Intelligence*

In 1904 Whipple (163) wrote a brief commentary on studies such as those of Bagley (10), Gilbert (59), and Wissler (167), the findings of which were assumed to indicate a relation between reaction time and general intelligence in children. He presented no experimental data, but pointed out the influence of varying conditions on reaction times secured from adult subjects in careful laboratory experiments. He suggested that positive relations found between reaction time and mental ability might be a reflection of better grasp of directions and closer following of instructions on the part of more able children.

Lemmon (84) obtained scores on simple and discriminative reaction time to visual stimuli for about 100 Columbia University undergraduates and related them to measures of intelligence, memory, and learning. The correlations of interest to us, those between reaction time and scores in the Thorndike *Intelligence Examination for High School Graduates*, ranged from .01 to .17, increasing in magnitude with each of five levels of difficulty in the discriminative reactions. The learning scores showed a similar trend in correlations with reaction time, the coefficients ranging from .11 to .35. Intercorrelations among the various reaction times varied from .54 between simple reaction times and times for the least complex discriminative reaction, to .27 between simple and more complex reactions. Reliabilities for the reaction time tests ranged from .88 to .95; for the learning tests, .83 to .95. The reliability of the Thorndike examination was given as .85 but was not obtained for this group. Lemmon concluded that although his correlations were all too low for predictive purposes they were consistent enough to indicate genuine trends, and he postulated speed of conduction in the nervous system as a possible factor responsible for the relationships revealed.

Peak and Boring (106) administered two forms each of the Army Alpha and Otis tests to five college students, examining the subjects individually and timing each separate item with a stop-watch. They controlled the factor of accuracy by taking as each person's time score the average number of seconds he spent on items which all five did correctly. This measure, obtained for Alpha and Otis items separately, was then correlated with simple reaction times and scores on Alpha and Otis obtained under conditions of standard and unlimited time. Rank order correlations between simple reaction time and average time on Alpha and Otis were .70 and .90; all other correlations between the measures of speed and test scores under standard time were .90 or unity. The authors concluded from these correlations that individual differences in intelligence[7] were due in the main to speed of reaction. They ignored completely their own correlations between the various speed measures and scores on Otis and Alpha obtained under unlimited time conditions. These ranged from —.20 to .10, and could be used as evidence leading to a completely different interpretation of the results. It need hardly be pointed out that conclusions based on data from five subjects are precarious.

The work of Farnsworth, Seashore, and Tinker (39) is in a sense a logical sequel to that of Peak and Boring. The number of subjects was extended to 34 college students, serial action measures were added to simple reaction time, and scores in the Ohio State University college entrance examination and Thorndike intelligence test were available as well as those on Army Alpha and the Otis advanced examination. Again time was taken on individual items of Alpha and Otis, but on this occasion through telegraphic recording on an ink polygraph. These authors found that attempts to hold accuracy constant by considering only items done correctly by all resulted in scores based on too few items to be a reliable measure. Hence the total group was divided into five subgroups and within each of the smaller samples time taken to complete the tasks done correctly by all was correlated by the rank order method with simple reaction time. The correlations ranged from —.66 to .90. This finding the authors took as evidence against the existence of a general speed variable in the two types of speed measures. Using the entire group of 34 subjects, Farnsworth, Seashore, and Tinker found a correlation between simple reaction and serial action times of .15. The four intelligence tests, given under standard time conditions, showed correlations with simple reaction time of from —.16 to —.24 ($PE = .11$) and with serial action time of from .14 (Ohio) to .53 (Alpha). When time on three of the standard tests was reduced to half or quarter time, relations with serial reaction time were a little more marked for Ohio and Otis but showed no difference in Army Alpha. Coefficients of correlation between unlimited scores and serial action were again low for the Ohio and Otis (.07 and .10) but higher for Alpha (.36±.10). Reliabilities of the simple reaction and serial action scores were reported as .72 and .91. The authors concluded that what the different intelligence tests measure varies and that the Alpha tends to become a test of serial action whereas the Ohio and Thorndike seem to be tests of content in which the speed factor is unimportant.

---

[7] Peak and Boring, on the basis of these results, define intelligence as mental power or work done against time. The difference between this concept and that of altitude, or level of difficulty independent of speed of performance, must be kept in mind. It is to avoid confusion of the two notions that contrast is made in this study between speed and altitude rather than between speed and power.

Philip (111) reported a study on reaction times of 311 Canadian children aged 9 to 16 in grades three to eight. When chronological age was held constant, correlations between reaction time and mental age measured by the Otis *Self-Administering Tests of Mental Ability* ranged from —.12 to .16 for girls, and from —.12 to .10 for boys, all correlations being statistically insignificant.

The trend of these results is toward the conclusion that the correlation between simple reaction time and mental ability as measured by standard intelligence tests is negligible. When simple reaction time is correlated with time taken to perform mental test items, with accuracy constant, the coefficients obtained from small groups may range from —.66 to .90, showing no consistent relationship between speed in simple reaction time and speed on mental tasks well within the ability of the subjects tested. Speed of discriminative reactions and rate of serial action seem to show higher correlations with each other or with measures of intelligence as their complexity increases, but the coefficients are too low for predictive purposes, and comparisons between these speed measures and scores on tests of mental ability seem to vary with the nature of the intelligence examination used.

3. *Speed of Simple Motor and Mental Functions as Related to Intelligence*

Hollingworth and Monahan (69) compared tapping rates of 50 children of superior intelligence with those of 50 matched with the first group as to age, sex, and race but unselected as to *IQ*. The average intelligence quotient (Stanford-Binet) of the experimental group was 151, and these children proved to be superior to the members of the control group in speed of movement in tapping. It was suggested that since an electrical counter was used to record tapping, this difference might be indicative merely of superior ability of the bright children to acquire a technique for making each tap effective.

Garrison (54) used 60 college students as subjects for an investigation of some simple speed functions and their relation to scores on speed and altitude tests of intelligence. His rate measures included tests of cancellation, card sorting and card distribution, substitution, tallying, and speed of vocal responses. He used as a speed test of intelligence the Otis *Self-Administering Tests of Mental Ability* and selected as a measure of altitude the analogies devised by Thurstone for the college entrance test of the American Council on Education. Correlations between scores on the various speed measures and the tests of intelligence ranged from —.22 to .34 with probable errors of .08 and .09. Correlations between the speed tests and the altitude measures were, contrary to expectation, a little higher than those with Otis scores. Intercorrelations among the speed tests ranged from —.08 to .77 with a tendency for the higher ones to obtain between measures showing similarity in content. This study gives no support either to the notion of a consistent relation between speed and altitude or to the concept of a general speed factor.

Walters (160) obtained a variety of measures on 165 New York public school children from the high sixth and low seventh grades. Mental ratings were obtained from the National, Otis, Pintner *Non-Language* and Pressey *Cross-Out* tests and from the Trabue and Stockbridge *Mentimeter*. Reading scores were obtained on *Stanford Achievement*, *IER*, Burgess, Monroe, and Thorndike reading examinations. Rate measures included two cancellation tests, a Courtis arithmetic test, and scores on the National and *IER* examinations. Additional speed scores were obtained by calculating rate of work on the National and Otis from the amount of work attempted in half time and by recording the added time taken by the children to finish the mental tests under unlimited time conditions. Mental age obtained on the *Stanford Revision of the Binet Scale* was taken as one measure of altitude, and a further measure was derived from a composite of *MA* and school success as measured by marks, teachers' ratings, and scores on standard achievement tests. Intercorrelations of the various rate measures ranged from .17 to .69. Correlations between the various separate rate and altitude measures, and between composites of each, ranged from —.0015 to .45. Those between scores on the group tests involving language and Stanford-Binet mental age were .53 to .68 under standard time conditions and were not altered by allowing unlimited time for the former. When a composite criterion of altitude was used, however, scores under unlimited time gave somewhat higher correlations with ability than did scores under standard time and ranged from .59 to .79. It was the author's conclusion that general intelligence could not be reliably predicted from measures of rate of performance.

Beck (11) reported a study on 30 university men and women from whom he obtained scores on the Thorndike college entrance examination and measures of speed of simple and serial reaction and reading. Reliabilities of the various measures ranged from .85 to .96, intercorrelations among the different speed scores from —.09 to .23. Correlations between measures of rate of performance and percentile score on the Thorndike varied from .02 (simple reaction speed) to .32 (reading rate). It will be seen that these results are consistent with those reported by other investigators who used reaction times. Beck concluded that there is little relationship between various kinds of speed and that rate of work is not the principal factor in success on the Thorndike test.

Another study which utilized speed scores from simple and complex motor and mental measures is that of McFarland (91). Timing the items individually, he administered tests, varying in degree of difficulty from simple auditory reaction time to insight problems, to three groups composed of four, five, and 34 college students respectively. He interpreted correlations among the various measures of from .00 to .88 to indicate that speed is an individual trait characterizing mental behavior. He reported no reliability coefficients for his measures and seems to have gone beyond the implications of his data in construing them to show a general speed factor.

Intercorrelations among complex functions here again appear to be higher than those between simpler ones. McFarland has, like Rounds, no measure of what is commonly thought of as altitude or level of intelligence. He appears to believe, along with Peak and Boring, that intelligence is defined in terms of speed of response and that therefore his results show the significance of speed in mental ability.

In the study by Dowd (35) mentioned earlier in another connection, speed scores in cancellation, hand movement, and writing tests yielded correlations of less than .10 with scores on the Otis advanced examination. The study was carried out with 165 sixth graders as subjects, and reliabilities of the various speed measures were .90 or above, that of the Otis test, .87. High intercorrelations among the rate scores were between tests similar in content.

In general the results of these studies indicate but slight and insignificant relationships between a variety of speed scores and different measures of intelligence. The findings of studies on simpler measures of rate of performance are borne out in that there is evidence of specificity of speed of response, higher intercorrelations being found among complex processes, where there is presumably more opportunity for overlapping of function, and among those similar in nature.

4. *Speed on Simple Mental Test Material as Related to Altitude*

Highsmith, Hunsicker, Clark, and Tryon and Jones have all attempted to measure speed and altitude in the same function by use of the same sort of material, varying the levels of difficulty. Highsmith (65) took as his criterion of altitude the *IQ* obtained from the *Stanford Revision of the Binet Scale.* The *National Intelligence Test* was administered as one in which score was influenced both by speed and by quality of performance. Rate tests consisted of linguistic and non-linguistic material from standard tests well within the level of ability of the group. The subjects were 87 boys and girls from the fifth, sixth, and seventh grades of the Peabody Demonstration School. The averages of correlations in the separate grades between altitude as measured by *IQ* and speed as measured by score, in reduced time, on easy intelligence test items were .14 for non-linguistic material, .31 for linguistic. *N. I. T.* point score correlated to the extent of .79 with linguistic rate, .39 with *IQ*, .30 with non-linguistic rate. The author concluded that linguistic rate tests were nearly as good an indication of intelligence as were the National tests but that neither was adequate. Reliabilities for the rate tests ranged from .53 to .80 and Highsmith pointed out that they were obtained on tests for which the time allowed was very short. The correlation between score on the *N. I. T.* and linguistic rate is the only one which is statistically significant in the usual meaning of the expression.

Hunsicker (70) worked with 163 subjects ranging in educational status from the fifth grade to graduate standing in college. Records for each person were taken individually and consisted of scores on a graded series of arithmetic and completion problems. The score for rate of work was taken from time to complete the first two pages of easy items. Altitude was indicated by the highest level of difficulty at which a person could do 50 per cent of the items correctly. Reliabilities for the various subgroups taken separately ranged from .75 to .95 for rate measures and from .75 to .90 for level. Coefficients, corrected for attenuation, between rate and altitude in both completion and arithmetic varied from .39 to .61, which the author took to be evidence of a consistent positive relationship between ability and rate of work. Considering ability in completion and arithmetic to be due to general intelligence, she further concluded that these correlations were indicative of the relationship between speed and level of general mental ability. This seems a rather narrow interpretation of intelligence. Intercorrelations between rate in arithmetic and completion ranged from .57 to .81 for the various groups and were taken as evidence of a general speed trait. In view of the findings as to relationships in a wider sample of speed measures, it would be possible to attribute these results to similarities of complexity or content in the measures, or to both.

Clark (27) used material similar to Hunsicker's in obtaining speed and level scores for 180 school children of grades seven to twelve, but time was taken by stop-watch for the individual items rather than for sets of items. He obtained one range measure by use of the *IER* information test and another by combining this with scores on the easier levels of the arithmetic and completion items. Level, speed, and range measures were all correlated with Stanford-Binet mental age and with scores on the Terman *Group Test* and the Otis *Self-Administering Group Test*. Coefficients, corrected for attenuation, were of the order of .54 between speed and the various intelligence measures, .70 between level and intelligence, .69 to

.77 between range and mental test score. Reliabilities of .79 to .88 were quoted for the different speed, level, and range measures, and those of the standard intelligence tests were taken to be unity. It seems likely that part of the correlation shown between speed and intelligence could be accounted for by the very wide grade range represented in the sample. Dowd (35) has pointed out that a further factor leading to these relatively high correlations between rate measures and intelligence tests is the similarity in content of the two.

Tryon and Jones (154), using a different technique, investigated the relation between rate of work as indicated by success on simple completion tests and altitude as measured by the completion items of levels *M* to *Q* in the *CAVD* examination. The 116 college students who served as subjects read simple narrative and descriptive material at four different exposure rates. The various sections of the test were presented on a motion picture screen, and the subjects were required to fill in a completion test on each unit after exposure. It was argued by the authors that, although pure speed of mental processes was not measured by this means, the absence of a relationship between speed and altitude would be indicated if correlations between level of difficulty and speed scores remained the same over a wide range of exposure rates. Reliability coefficients for the speed tests ranged from .84 to .92 for the four exposure times; the reliability of the altitude measure was .94. Correlations between altitude and the various speed scores, when corrected for attenuation, were of the order of .50 and did not differ significantly from each other. Intercorrelations between speed scores were in the neighborhood of .80 and seemed to indicate that individual differences as measured by this task were not much influenced by the exposure rate factor. The authors concluded that their results lend no support to the notion that level or altitude is dependent upon the rate of work.

These studies have shown results leading to arguments both for and against a relationship between rate of performance on simple material of the mental test type and various measures of altitude. It is noteworthy that conflicting conclusions have been indicated where one might expect a fair degree of relationship to exist on account of similarity of content in the measures of speed and level. The weight of evidence appears to be on the negative side, since the two studies supporting the notion most strongly can be criticized on the score of narrowness of the function measured (70) or wide age range of the sample used (27). Highsmith's correlation of .79 between a rate score and level as measured by a standard intelligence test points to a possible criticism of the work of Clark and all others who use as their criterion of altitude a test score obtained under rigorous time limits.

5. *Speed and Intelligence as Measured by the Same Test*

The method used by the authors whose work is discussed under this head is similar to that employed in the studies described in Section 4. May, Ruch, and Koerth, and F. S. Freeman have all compared scores made on standard intelligence tests under regular time conditions with those earned in unlimited time, and to a certain extent this amounts to comparing scores on easy items with scores on both easy and more difficult items. However, certain differences in interpretation of the results, together with the fact that the measures compared have been derived from the same instrument in each case, make it convenient to discuss these three studies separately.

May (89) reported the results of the application of the Army Alpha to 510 army men with standard and double time allowances. The correlation between scores obtained under the two conditions of administration was .96, which led May to assert that extension of time on the test would not materially alter the rating of a given individual. An analysis of the per cent of persons at various intelligence levels who gained in score under double time showed that fewer low scorers gained and the implication was that the Army Alpha is more nearly a "power"[8] test for such individuals. May concluded that the Army Alpha is essentially a speed test but drew no inferences as to the relation between speed and level.

Ruch and Koerth (121), as a check on the army study, gave the Alpha to 122 college freshmen, 70 from the lowest decile in the class, 52 from the highest, as determined by three other standardized tests. Scores were available for standard, double, and unlimited time and showed results similar to those obtained by May. The correlation between single and double time was .97, that between single and unlimited time, .94. Ruch (120) later used the same technique in testing 150 seventh and eighth graders with the Terman *Group Test* and found a correlation between standard and unlimited time scores of .96. Again the conclusion was simply that the speed factor presumably present in the test under standard conditions did not interfere to any great extent with the ranking of the subjects.

---

[8] May, Ruch, and Koerth, and F. S. Freeman apparently use this term as equivalent to what we have been referring to as altitude.

In a series of three studies, F. S. Freeman (48, 49, 50), examined 100 children from grades four to seven, 149 pupils from grades six, eight, and eleven, and 177 college students. In each case he administered intelligence tests under standard time conditions and then repeated them without time limits at intervals of from one day to three months. He proceeded on the assumption that high correlation between scores in single and double time on a time-limit test might indicate either that the test was primarily one of power or that speed was still entering into the score under double time conditions. He reasoned that if correlations were high between scores on the same test in standard time and with all time limits removed then the examination would be one of power, that is, not affected by time limits. Correlations between scores obtained under the two different conditions of timing varied, for the separate studies, from .58 to .93. This led the author to state that speed is of minor importance in determining success on intelligence tests for most subjects. Because the correlations failed to reach unity, Freeman concluded that there was a small group adversely affected by time limits. Tryon (153) has pointed out, however, that the correlations were uncorrected for errors of measurement, hence that evidence for the existence of this group penalized by time limits is not conclusive.

The studies of May and of Ruch and Koerth show little beside the fact that imposition of a time limit does not interfere much with the relative rankings of individuals on the Army Alpha test. Perhaps F. S. Freeman's most significant contribution is his statement that correlations such as they obtained are not conclusive evidence of whether speed or "power" is measured by the tests or what the relationship may be between the two. Jones and Tryon (154) have further pointed out that the so-called speed measures obtained under standard timing conditions also contain an altitude element.

A few scattered studies remain to be considered, the most important of which is that of Bernstein (12). This author devised tests of completions, directions, analogies, concomitants, and moral classifications for administration under "leisure" and "haste" conditions. These conditions were arranged by presenting varying numbers of items to be completed within a given time. Scores on the different leisure tests were pooled and correlated with teachers' estimates of intelligence and slowness, and the same procedure was followed with the haste tests. Intelligence ratings correlated .556 with scores on the leisure tests and the same (.558) with scores on the haste tests. Scores on leisure and on haste tests correlated with ratings on slowness gave coefficients of —.45 and —.37. The relationship between the haste and leisure scores was as high as the average intercorrelations among the leisure items and the haste items. Correlations between a measure of slowness (the difference between leisure and haste scores) and ratings for intelligence and slowness were not significantly different from zero. Bernstein's data were obtained on one group of 70 English school boys about 14 years old and on another similar group of 60. The results were taken as evidence against the existence of a speed ability apart from general intelligence and are important because they were used by Spearman in this connection. Sutherland (137) has since pointed out that the time limits for Bernstein's leisure tests were so short that it is doubtful that leisure conditions were produced; hence, correlations between leisure and haste tests and intelligence were really correlations between intelligence ratings and two different speed scores.

Slater (130) has presented evidence in support of Bernstein's conclusion. Scores were obtained on an adaptation of the *CAVD* and on five separate non-verbal mental tests administered with a time limit for 226 school children 13 to 14 years of age. The subjects noted the time of beginning each item on the *CAVD* from a specially constructed timing device set up at the front of the room. The average amount of time taken by the group to solve problems correctly was found for *C*, *A*, *V*, and *D* separately, and individual speed rates were obtained in terms of deviations from these averages. Correlations were obtained between speed and level scores on the separate sections of the *CAVD* and between these and scores on the non-verbal tests for various sub-samples of the total group. Correlations in the total array ranged from —.26 to .68, those between level scores from .03 to .67, those between speed rates from .46 to .68, and those between level and speed from —.26 to .48. Speed rates tended to be consistent for a given child but to show no very close relationship with intelligence as measured by verbal or non-verbal material given with or without time limits. The classification test, one of the non-verbal measures, showed the highest correlation with speed rate, .48, yet did not require quicker response than other tests. Another non-verbal item, the code test, which in the judgment of the author demands the greatest quickness of judgment under pressure, showed low correlation with speed measures, —.26 to .29. These findings led Slater to conclude that conditions of pressure do not invalidate mental test results by introducing a special factor of speed.

With 193 college students as subjects Saum (123) correlated scores obtained from the *Ohio State Psychological Test,* Form 14, taken under time limits, with scores from Forms 13, 15, and 16 taken as a work limit test. The coefficients ranged from .71 to .77 for several subgroups with an average correlation of .74. A scholarship measure, based on a point-hour ratio, correlated with time limit scores and work limit scores to the extent of .60 and .65. The author interpreted her findings as indicating that time is a negligible factor in college level intelligence tests.

Chapman (26) reported an average rank difference correlation of .28 between "speed of thought" and "success" in a test of word building. Speed of thought was measured by the average time per word, calculated from total time for the first 10 words; the measure of success was the total number of words formed from the six letters which were given. When all subjects were required to spend 15 minutes at the task the average correlation between speed and success rose to .63. A total of 135 eighth grade children worked in four groups and correlations were calculated for each group separately. The evidence of the operation of a factor of persistence which influenced the results, the unreliability of the speed measure (about .49), and the narrowness of the function tested make the results of little value for the present problem.

F. N. Freeman (45) reported a correlation of .00 between score on the Burt reasoning test and time taken to complete the test. Little information is given about the sample save that it consisted of high school students. As Hunsicker (70) points out, Freeman was in error in assuming that rate of work is measured by the time taken to produce a maximum score on a test of altitude.

This survey of the literature has revealed disagreement among theorists as to the nature of speed and its relation to intelligence, and it has brought to light conflicting findings from experimental investigations of the problem. In order to attack the question of the effect of time limits on intelligence test performance it will be desirable to make further study of the relationship between rate of work and level of ability.

III. THE EXPERIMENT

A. The Problem

The effect of rate of work on test performance may be more clearly indicated if some of the issues raised by earlier workers in the field are further pursued.

As has been mentioned before, the objection to group tests of intelligence because of their emphasis on speed of performance is based on the assumption of a negligible relationship between rate and altitude or level of mental ability. It has been shown, however, that theorists are not agreed as to the nature of this relationship, Spearman holding that speed is an essential part of intelligence, Thorndike arguing for the notion that the two are separate aspects of mental ability, usually, but not always, found in positive relation to each other. The findings of those who have investigated the interrelationships among speeds on a variety of different tasks point to the conclusion that a general speed ability characterizing the work of a given individual is by no means demonstrated. This fact, together with the finding of correlations of widely varying magnitudes between rate of performance on simple mental and motor tests and measures of intelligence suggests the futility of investigating the relationship between speed and altitude when the measures of rate and level are obtained on different tasks.

The work done to date on the relationship between speed and altitude when both are measured by performance on intelligence tests needs to be extended because of conflicting findings. It seems likely that some of the results which indicate a fairly high relationship between rate of performance and altitude are due to inadequate measures of speed and level (50, 89, 121), heterogeneity of the sample (27), or narrowness of the function tested (70). Of the studies which indicate absence of such a relationship, one employed tests of differing content for the speed and level measures (65), the other was confined to completion items only (154).

The aim of this study is to discover what relationship, if any, exists between rate of successful work on a standard intelligence test and level as indicated by the same examination as well as by other measures of altitude. Also, it will be pertinent to inquire whether there is any difference in the reliability or validity of such a test when it is given under standard and unlimited time conditions. The possible effect of age on the results will be limited by confining the experiment to low seventh grade children.[9] The Terman *Group Test* will be employed for determination of the speed and altitude scores because of its wide use and its applicability to children of this grade. Items of the test will be timed individually so that rate of successful work may be indicated by the average amount of time the child spends on items he does correctly, and altitude on the test will be measured by score obtained under unlimited time conditions. Justification of these procedures will be presented when they are more fully described. Results for the two sexes will be treated separately in order that any sex differences present may be revealed.

B. THE SAMPLE

The 200 boys and girls who comprised the main sample of the present study were drawn from the low seventh grade of the Claremont Junior High School in Oakland, California. This group, with four exceptions, was the one which had been selected a year earlier for the intensive study on adolescence being carried out by the Institute of Child Welfare at the University of California. The bases of selection had been prospective attendance at the Claremont School, probable permanence of residence, and willingness of parents to coöperate with the Institute staff in the long term project.

The group covered a grade range of one school year, half of the children being in the lower grade at a given time, half in the higher. Table 1 shows the distribution of the sample according to age and sex for each grade separately. The age in low seventh grade is presented, which means that the figures are representative of the higher grade in the spring of 1933, of the lower grade in the fall of the same year. The date taken for calculation of exact age was that of the first individual test administered.[10]

Further descriptions of the sample, drawn from the school records and the files of the Institute of Child Welfare, are presented in Tables 2 and 4. Table 2 shows the distribution of the children according to the intelligence quotient derived from the Kuhlmann-Anderson test administered to each half of the group in the high

---

[9] This precaution seems particularly desirable in view of the possible change with age of the speed-power relationship as indicated by Miles (96) and others (18, 29, 86, 94).

[10] See below under procedure.

TABLE 1

DISTRIBUTION OF CHRONOLOGICAL AGES FOR EXPERIMENTAL GROUP AT TIME OF FIRST INDIVIDUAL TEST

Age in years and months	Boys Higher grade	 Lower grade	Girls Higher grade	 Lower grade
13^6—13^8	1			
13^3—13^5	0			4
13^0—13^2	2		2	3
12^9—12^{11}	4	3	4	3
12^6—12^8	8	5	7	7
12^3—12^5	9	12	9	14
12^0—12^2	15	11	12	12
11^9—11^{11}	5	8	7	6
11^6—11^8	4	5	4	4
11^3—11^5	2	0	4	1
11^0—11^2	1	1		1
Number	51	45	49	55
Mean age	12.25 yrs.	12.17 yrs.	12.17 yrs.	12.33 yrs.
SD	.50 yr.	.33 yr.	.42 yr.	.50 yr.
Mean age of both grades combined	12.17 yrs.		12.25 yrs.	
SD of both grades combined	.42 yr.		.50 yr.	

TABLE 2

DISTRIBUTION OF INTELLIGENCE QUOTIENTS FOR EXPERIMENTAL GROUP

Kuhlmann-Anderson IQ	Boys Higher grade	 Lower grade	Girls Higher grade	 Lower grade
130—134		3	1	1
125—129	1	1	0	1
120—124	1	1	1	1
115—119	5	9	2	2
110—114	3	8	5	8
105—109	9	8	11	15
100—104	14	8	10	11
95— 99	7	4	6	5
90— 94	7	3	8	4
85— 89	2		3	2
80— 84	1		0	2
75— 79	0		1	1
70— 74	1		1	
Number	51	45	49	53*
Mean *IQ*	102.2	109.6	101.5	103.9
SD	10.2	10.1	10.8	10.4
Mean *IQ* of both grades combined	105.6		102.7	
SD of both grades combined	10.8		10.7	

*No test scores available for two subjects.

TABLE 3

SIGNIFICANCE OF THE DIFFERENCES BETWEEN THE MEAN INTELLIGENCE QUOTIENTS FOR THE HIGHER AND LOWER GRADES

	Boys	Girls
Difference in *IQ*	7.4	2.4
$PE_{diff.}$	1.40	1.42
D/PE_{diff}	5.29	1.69

TABLE 4

DISTRIBUTION OF AVERAGE BARR SCALE RATINGS FOR EXPERIMENTAL GROUP

Rating	Number of cases
15.0—15.9	12
14.0—14.9	2
13.0—13.9	17
12.0—12.9	22
11.0—11.9	36
10.0—10.9	28
9.0— 9.9	27
8.0— 8.9	11
7.0— 7.9	14
6.0— 6.9	2
5.0— 5.9	2
4.0— 4.9	2
3.0— 3.9	2
Number	177
Mean	10.90
SD	2.43

sixth grade. This examination was part of the regular testing program of the Oakland Public Schools and was given approximately six months before the individual tests of the present investigation. It will be seen that in respect to intelligence this group is a normal one in the usual meaning of the term (139, p. 78).

All subjects were American-born whites and the majority were children of American-born parents. In only six per cent of the cases were both parents born in a non-English-speaking country and in but three families was there no English spoken at home. It seems unlikely that the children from the homes in which only a foreign language was spoken would be handicapped in any way seriously affecting the results of this experiment since they had all been using English in school for at least six years.

Four independent Barr scale[11] ratings were available for 177 of the cases. Terman (142, p. 71) gives the mean rating of the adult males of the general population as 8.88 or 7.92 with standard deviations of 3.24 and 3.38. The distribution on Table 4 is of the average of the four ratings and shows that neither extreme is represented in this sample, but that the mean rating is somewhat superior to that of the general population.

In the presentation of results, most of the material for higher and lower grades will be combined for each sex. Table 1 has shown that there is very little difference in mean ages for the two sections of boys and of girls. Table 3 reveals a significant difference in intelligence quotient between the higher and lower grades of boys, but this probably means only that neither smaller group alone is truly representative of the general population of low seventh grade boys in this school. Since neither the comparisons presented in these tables nor the circumstances of selection of the sample make it seem likely that there are systematic factors operating to differentiate the two groups, the combination of data obtained on two successive low seventh grades should lead to more dependable results.

C. INSTRUMENTS AND APPARATUS

1. *The Test*

As was mentioned previously, the Terman *Group Test of Mental Ability* was used for the present investigation. This test is intended for grades seven to twelve and is widely employed, two facts which recommend it for our purpose. Since administration of the Terman test at two-year intervals was part of the program of the study on adolescence, its use insured the possibility of comparisons between early and later scores on the same examination and served to further the plan of the Institute investigation. The manner in which Forms *A* and *B* were used, the changing of instructions, and the omission of certain sub-tests will be discussed in the section on procedure. For full description of the test the reader is referred to the published examination and to the manual of directions (141).

---

[11] The Barr scale presents for 121 occupations numerical values ranging from zero to 20.71 according to the intelligence demanded for successful performance of each type of work. See (142, pp. 66-71).

2. *The Rating Scale*

On the basis of observation during the individual examinations, ratings were made on 19 behavior traits for 181 children of the total experimental sample. The seven-point scale employed was adapted from one in use at the Institute of Child Welfare for ratings on pre-school children during an intelligence test. Miss Judith Chaffey, who was administering the Stanford Revision to the adolescent group and making use of the rating scale at the time of this study, collaborated with the writer in defining the points on the scale. A copy of the scale and the description of the various ratings for each trait are to be found in Appendix *A*. It will be noted that in all but two cases each of the seven points is defined for every trait. The descriptions were not considered either exhaustive or final but merely suggestive of the various scale levels.

3. *The Short Sample Record*

It was thought that some record of the overt behavior of the child, more detailed and specific than that available through the ratings, might be of value in interpretation of certain of the test results. Consequently, during the early part of the testing program, an effort was made to devise a check list of items of behavior the presence or absence of which in a given short period of time might be noted. It was soon apparent that a list comprehensive enough to include all items which might appear would be too long to be practical in the test situation set up for this investigation. Accordingly a sheet of cross-section paper was ruled to provide one square for each 30 seconds of actual testing time. On this sheet each behavior item was written in as it appeared for a particular child, and a tally mark was made in the appropriate square for each 30-second interval throughout the examination in which that type of behavior occurred. In this manner a rough record was kept of body movements, vocalizations, attempts to establish contact with the examiner, and expressions of boredom or fatigue.

4. *Bristol Strip Recorder*

The Bristol Strip Recorder[12] is an instrument adapted to measuring in seconds or half-seconds the time spent on each of a series of specific activities, such as answering items of an intelligence test. It is equipped with multiple writing pens which make a record in ink on a specially printed strip of lined paper. Uniform rate of movement of the recording paper is assured by the fact that there are two motors, one of which drives the mechanism while the other controls the rewind rolls. The power from the drive motor is applied to a roller over which the paper moves as it passes from stock roll to rewind roll. The gears used in the present experiment gave a rate of movement of three inches per minute. The writing pens were electrically connected with a push button held in the experimenter's hand; this made it possible to keep the recording device itself out of sight of the subject and allowed the examiner to give full attention to the child's performance. The technique of reading the records will be discussed in another section.

---

[12] This instrument is manufactured by the Bristol Company, Waterbury, Connecticut.

D. Procedure

In order to obtain a measure of time spent on each item answered correctly, each member of the experimental sample was examined individually on the Terman *Group Test of Mental Ability*. For administrative reasons, the total time spent on this test was reduced by omitting three sub-tests which have a relatively low item reliability (Tests 3, 6, and 8, consisting of items which present only two choices from which the correct answer is to be selected). During this examination standard timing for each portion of the test was used and the number of seconds spent on each item was indicated on the Bristol recorder, completion of work on a given item being noted when the subject finished the marking necessary to indicate his answer.

The timing apparatus was placed on a low bench beside the examiner, behind the desk on which the subject wrote and out of his sight. The writing pens made a slight clicking noise as they moved and 11 out of the 200 children gave evidence of noticing it. Five of them merely glanced up at the first sound and then went on with their work. Six of them asked about it and were told that a mark was being made whenever they finished an item. None of the children revealed that he had guessed timing of the items to be the purpose of the record, and speed of work was never mentioned in the few conversations on the subject. The subjects did not appear to be in the least disturbed by the noise of the timer and it is probably safe to assume that its influence was not great on the performance of children as thoroughly accustomed to testing situations as was this group.

Each child in the group also took the alternative form of the Terman *Group Test*, including all sub-tests, under unlimited time conditions. On this occasion the subject's progress at the end of the standard time allowance was noted, and total time spent on each sub-test was taken by means of a stop-watch.

The instructions given in the test manual (141) were altered somewhat to fit the special needs of this investigation. It was necessary to add certain directions which would make recording of time spent on individual items more accurate, and it also seemed desirable to eliminate those parts of the standard instructions which are useful only when the test is being administered to a group. When the shortened examination was being given under standard time conditions it was necessary to warn the subjects that they must stop at the signal. On the other hand, under conditions of unlimited time the child was to be allowed to attack each item. In order to guard as much as possible against a hurried attitude on the one hand and dawdling on the other, the command *"Go as fast as you can without hurrying"* was substituted for *"Both speed and accuracy will count toward your score"* in both testing situations. Every effort was made to deliver this statement in a manner which would bring about the desired attitude. The full test of revised directions for all testing conditions is to be found in Appendix *B*.

As will be noted by reference to Appendix *B*, in the unlimited time test the child was not told that he could work as long as he wished, but time was called only when he indicated that he had done as much as possible. The question was then asked, *"Are you sure you have done all you could?"*, and the subject was urged to go on if there was any doubt in the examiner's mind as to whether every item had been attempted. The only exception to this procedure was in the case of Sub-test 5. This is an arithmetic test, the most difficult items of which were beyond the level of ability of any child in the group. Time on this sub-test was called at eight minutes, or double the standard allowance, when it became obvious that no child would increase his score if given more time.

Immediately after each examination, the rating schedule was filled in and incidental notes were made on testing conditions and on the subject's comments or other behavior. Short sample records were also made during the complete examination, and the subjects were asked the following questions: (*a*) *How do you like this kind of a test?* (*b*) *Would you rather take it alone this way or with the whole class? Why?*

Approximately half of the total group took the shortened test first and half the complete test. The sub-groups thus formed were further divided so that half of each had Form *A* first and half Form *B*. This procedure was adopted as a means of controlling possible effects of position in series in relation to the timing procedure and the form of the test. All scoring of the tests, and subsequent computations, have been checked for possible errors.

The individual examinations were administered in the period between February 8 and May 24, 1933, for the higher grade, and between September 5 and November 21 of the same year for the lower grade. In the spring term 60 per cent of the children had the complete test at school and the shortened one at the Institute of Child Welfare. The other 40 per cent of this group, and all of the children tested in the fall, took both examinations at the school. In all cases, the tests were given in rooms scheduled for this purpose and without interference from other activities.

Children were excused from their classes between 8:20 and 11:40 A. M. and between 12:30 and 3:20 P. M. to take examinations given at the school, whereas those who were seen at the Institute were tested between three and five o'clock in the afternoon. It was assumed that differences in time and place of testing, within the limits here described, would not materially affect results, either in comparisons of one child with another or between the two tests of the same subject.

The intervals between the two examinations for each subject ranged from seven to 27 days, with all but one subject taking the second test within three weeks of the first. Out of the total of 400 individual tests, 25 or 6.25 per cent were administered in two parts because of unavoidable interruptions. In 21 cases the children had used all of one class period in attacking the test under unlimited time conditions and could not miss the next class or had to be excused to go home. In two cases fire drill interrupted a test which could not be immediately resumed, and in the two remaining instances the timing apparatus broke down. Out of the 25 cases of divided tests, eleven were completed in another period on the same day, four the next day, and ten from two to seven days later due to absence of the child or delay in repairing of the Bristol recorder. There was no evidence that these disturbances interfered seriously with any child's performance, and the total number of cases so affected was probably too small to influence significantly the results for the whole group.

IV. RESULTS

A. THE DATA

1. *Terman Group Test Scores*

In presenting the mental test scores it was desirable to combine in one distribution measures obtained from the two forms of the Terman *Group Test* under two different methods of administration. In order to demonstrate that such combination of data was justified, Table 5 was prepared showing the average point score on each form of the examination when given as first and as second test and when administered under conditions of standard and unlimited time. For the sake of brevity these conditions are referred to as speed and non-speed conditions respectively. Average scores on seven sub-tests are presented since examinations under speed conditions were limited to this number. The eight sub-groups, formed by division of the total sample according to sex, time conditions of first test, and form of first test, have populations of from 21 to 28 subjects each.

TABLE 5

AVERAGE POINT SCORE ON SEVEN SUB-TESTS OF THE TERMAN GROUP TEST FOR VARIOUS SUB-GROUPS OF THE TOTAL EXPERIMENTAL SAMPLE

	Form: A	A	B	B	$\frac{A+B}{2}$	$\frac{A+B}{2}$	$\frac{A+B}{2}$
Sub-groups	Order: I	II	I	II	I	II	$\frac{I+II}{2}$
Boys: tested under speed conditions	68	78	67	75	67	77	72
Boys: tested under non-speed conditions	76	81	68	71	72	76	74
Girls: tested under speed conditions	62	62	52	71	57	66	62
Girls: tested under non-speed conditions	67	64	54	64	61	64	62
Total: tested under speed conditions	65	70	59	73	62	72	67
Total: tested under non-speed conditions	71	72	61	67	66	70	68
Total boys speed and non-speed	72	79	67	73	69	76	73
Total girls speed and non-speed	64	63	53	67	59	65	62

Briefly to summarize the material in Table 5, it may be said that the only consistent differences found are those not concerned with the method of administration of the test. Boys show consistently higher average scores regardless of timing conditions, and, as was to be expected, averages on second tests are in general somewhat higher than those on first tests, although this does not hold for girls tested on Form *A*. Form *B* appears to be more difficult than Form *A* when given as the *first* test for all groups except boys tested under speed conditions, but this is not consistently true when it is given as the second test. The probable errors of the largest differences found in the three sets of comparisons were calculated. These values, together with the differences between means and the critical ratios are shown in Table 6. It should be remembered that comparisons between first and second tests for the same form and same conditions of administration do not involve the same subjects.

TABLE 6

COMPARISON OF VARIOUS SUB-GROUPS OF THE TOTAL EXPERIMENTAL SAMPLE ON THE BASIS OF AVERAGE POINT SCORE ON SEVEN SUB-TESTS OF THE TERMAN GROUP TEST

Comparison	*Diff.*	PE_{diff}	*CR*
Boys *vs.* girls			
A II non-speed	17	2.96	5.74
I *vs.* II			
Girls B speed	19	2.99	6.02
A *vs.* B			
Girls I non-speed	13	2.98	4.36

Table 7 shows distributions of scores on first and second examinations,[13] regardless of form or method of administration, for the

---

[13] It will be noted that the totals fall short of the 96 for boys and 104 for girls making up the 200 children actually tested. In the case of one boy and two girls examined on the first day of the testing program, fewer than seven sub-tests, or a different selection, were given, hence these scores were omitted from the distributions for first tests. One girl was discovered to be a non-reader and both of her scores were thrown out. Two boys had low scores on the first examination because of failure to follow directions in one sub-test. A change of emphasis in the reading of instructions eliminated the difficulty of the second test and results on the first examination were discarded as not representative of the ability of these subjects. Two boys were given the same sub-test for the second examination as for the first and the second scores were omitted from the distribution as being probably too high. One boy had helped to score Terman tests for a student in practice teaching in the interval between his two examinations, and his second score was discarded as being probably influenced by this experience.

TABLE 7

DISTRIBUTION OF TERMAN GROUP TEST SCORES FOR TOTAL EXPERIMENTAL SAMPLE

Score	Boys		Girls	
	I$_7$*	II$_7$	I$_7$	II$_7$
115—124	3	1		
105—114	2	7		1
95—104	3	10	2	6
85— 94	14	13	8	7
75— 84	14	19	13	12
65— 74	16	15	13	29
55— 64	20	15	23	21
45— 54	13	9	16	11
35— 44	3	1	18	11
25— 34	4	1	7	5
15— 24	1	2	1	
Number	93	93	101	103
Mean	69.3	75.7	58.7	65.1
SD	20.5	20.4	17.7	18.0

*Symbols are to be interpreted as follows:
I, First testing
II, Second testing
$_7$, Seven sub-tests (1, 2, 4, 5, 7, 9, 10)

total experimental group. The average score is 6.4 points higher on the second test than on the first, for both boys and girls. Girls show consistently lower scores and smaller standard deviations than boys.

In Table 8 are presented the results of the complete tests (consisting of 10 subtests) given under conditions of unlimited time, and scored both for standard time limits and without time limits. Here separate distributions have been made for those subjects who had the complete test first and for those who had it as the second examination; both forms are represented in each distribution. The average scores of the two sets of subjects are almost identical in the case of boys and also of girls. A sex difference again appears, in that average scores and standard deviations of girls' distributions are lower than those for boys.

2. *Time Records*

The records of time spent on individual items on the shortened test were used as a basis for determining each child's rate of success-

TABLE 8

DISTRIBUTION OF TERMAN GROUP TEST SCORES FOR VARIOUS SUB-SAMPLES OF THE TOTAL EXPERIMENTAL GROUP

Score	Boys				Girls			
	TI*	TII	TUI	TUII	TI	TII	TUI	TUII
170—179				1				
160—169	2		2	2				
150—159	0	2	0	1			1	1
140—149	1	2	2	1			0	1
130—139	1	0	5	1	1	3	0	1
120—129	2	2	4	6	2	0	5	1
110—119	8	5	4	7	2	0	2	3
100—109	2	7	5	9	4	3	8	6
90— 99	6	6	3	6	7	3	6	8
80— 89	3	9	5	5	6	12	12	10
70— 79	6	9	3	5	10	10	7	7
60— 69	5	2	5	3	9	8	6	4
50— 59	3	3	2	1	8	4	3	4
40— 49	2	1	3	0	1	4	1	2
30— 39	2	0		1	1	2	3	1
20— 29		2		1	3			
Number	43	50	43	50	54	49	54	49
Mean score	90.5	90.7	99.2	101.9	76.2	76.3	86.0	86.7
SD	31.5	28.1	32.2	30.2	24.6	22.6	24.8	24.7

*Symbols are to be interpreted as follows:

T, Total test (10 subtests) scored on basis of standard time
TU, Total test scored on basis of unlimited time
I, First testing
II, Second testing

ful work. Before presenting the distributions of these scores it will be well to discuss their derivation and the reliability of the speed measure.

The mechanism of the Bristol recorder made it necessary to wind the drive motor and the take-up motor before each record was started. The rate of movement of the recording paper was timed by means of a stop-watch before and after each test, and the record was discarded if the check-up showed a change of rate in the instrument during the examination. Two tests were thrown out for this reason, and in four instances time records on one or more sub-tests had to be eliminated because the apparatus was out of order.

The recording paper was printed with lines one-tenth of an inch apart and as previously noted, travelled at a rate of three inches per minute. This meant that an interval of two seconds was represented in the space between two lines, and records could be read to the half-second by means of a ruler divided into fortieths of an inch. Three persons trained and supervised by the writer assisted in the work of measuring the distances between marks on the records. Two independent readings were made for each item, and one of the assistants re-read all records where there was a discrepancy in the measurements.[14]

As has already been indicated, it was thought desirable to eliminate from each child's rate score all time spent on items omitted or done incorrectly and to consider only time spent on items done correctly. It was also necessary to restrict the range of items to some extent, since the children who worked faster would reach the later and more difficult items in each sub-test and would hence be judged on a somewhat different basis than would the slower children. The method used by earlier investigators (39, 106, 130), that of using only those items which all subjects answered correctly, results in considerable reduction of the number of measures and consequent loss of reliability. It was therefore decided to use for each child's score the number of items he did correctly, but to limit the measures under consideration in each sub-test to those which were being attacked by at least 75 per cent of all subjects. It was assumed that within this range of difficulty, no serious error would be introduced by using slightly different selections of items for each individual's time score. On the other hand, more difficult items or those coming later in a subtest and attacked by less than 75 per cent of the total group were omitted from an individual's rate score even though he may have dealt with these items successfully. To summarize, then, a subject's speed or rate score was based upon (*a*) items which he answered correctly, and (*b*) items within a limited difficulty range, such that they were attempted by at least 75 per cent of the group.

---

[14] Except in rare cases of gross error, all differences in readings were of one half-second. Of a total of more than 10,000 item measurements 885 showed this discrepancy, and the third readings on these were examined in order to discover whether the person making the final measurement showed a tendency to read systematically high or low or to favor his own rather than another's reading. In 428 instances the final reading was the same as the higher of the first two, in 457 it was equal to the lower of the first two. In the 742 cases in which the same person made both second and third measurements, 411 of the third readings were the same as the second measurements, while 331 were equal to those of the first reader.

Certain irregularities in the records of a few of the children will be mentioned, although it is not likely that the speed scores were vitiated by their influence. With about seven per cent of the subjects, from two to four successive single items done correctly were timed together instead of separately. Since an average of time spent on all items done correctly was to be used, the timing of several items together should not have introduced an error and therefore these items were included in the scores. In approximately six and one-half per cent of the 200 cases, a child reversed the order of attack on the items at some point in the examination. For example, instead of doing items 1, 2, and 3 of a given sub-test in that order, the subject might have attempted 3 after 1 and then returned to 2. For two subjects out of the 200 it seemed obvious both from the time record and from observation of the child that work had been done on the omitted item both before and after the next one had been attacked. Therefore time measures on these items were not included in calculation of the rate score of the two children. In the case of the six and one-half per cent mentioned, however, the items omitted and returned to later seemed to have been overlooked at first, and time measurements on these were included among the ones which served as the basis of the speed score. A total of 19 single items so affected were distributed among the test records of the 13 children involved.

Reliability coefficients for the rate measures on each separate sub-test were calculated by correlating the total number of seconds spent on odd-numbered items with that spent on even-numbered items[15] and raising these correlations by use of the Spearman-Brown formula. The coefficients, which are presented in Table 9, were calculated for each grade separately. When the two grades are compared, the absence of any consistent differences in reliability seems to warrant combining these two grades into a single group.

In the final calculations, the items from Sub-test 5 were not used since so few of the children gave any correct answers to the arithmetic problems and the reliability of the time measurements on these items is consequently low. Table 10 shows the number of items attempted by 75 per cent of the group, and the possible time for each sub-test. Since only items done correctly were used for each child, most time scores on the separate sub-tests were based on fewer items and shorter performance times than are indicated in the table. The sub-test reliabilities obtained, averaging between .75 and .80 for the different grades and sexes, seem high in view of these limitations. Spearman-Brown reliability coefficients for rate score on Sub-tests 1, 2, 4, 7, 9, and 10 combined are:

---

[15] Odd-numbered items are, of course, the first, third, fifth, etc., done correctly, and not necessarily the odd-numbered ones according to the test booklet.

TABLE 9

SPEARMAN-BROWN RELIABILITY COEFFICIENTS OF TIME MEASUREMENTS ON ITEMS DONE CORRECTLY IN SEVEN SUB-TESTS OF THE TERMAN GROUP TEST

		Higher grade			Lower grade		
	Sub-test	No. of cases	Odd-even *r*	Rel. coeff.*	No. of cases	Odd-even *r*	Rel. coeff.
Boys							
	1	49	.789	.882	45	.746	.854
	2	46	.696	.821	45	.721	.838
	4	47	.821	.902	45	.832	.908
	5	34	.367	.536	37	.312	.475
	7	47	.643	.783	45	.405	.557
	9	47	.552	.711	45	.590	.742
	10	43	.545	.706	43	.723	.840
Girls							
	1	42	.532	.695	50	.478	.647
	2	44	.779	.876	53	.563	.720
	4	43	.820	.901	52	.829	.906
	5	22	.248	.398	25	.205	.340
	7	44	.575	.730	54	.572	.728
	9	44	.582	.736	54	.568	.725
	10	39	.440	.611	41	.650	.788

*Probable errors for the Spearman-Brown reliability coefficients, computed by Shen's formula (126), vary from .018 to .178. When Sub-test 5 is left out of consideration, the *PE*'s range from .018 to .085.

TABLE 10

TIME ALLOWED ON SEPARATE SUB-TESTS AND MAXIMUM NUMBER OF ITEMS AVAILABLE FOR CALCULATION OF RATE SCORE ON TERMAN TEST

		Maximum number of items			
		Boys		Girls	
Sub-test	No. of minutes allowed	Higher grade	Lower grade	Higher grade	Lower grade
1	2	16	16	14	13
2	2	9	9	10	8
4	3	17	18	17	16
5	4	5	6	6	5
7	2	12	13	12	13
9	3	18	18	18	18
10	4	10	11	10	10

Boys (both grades)	.947±.007
Girls (both grades)	.922±.010

The average amount of time spent on items done correctly on the different sub-tests was found for each child. Scores for the upper and lower grades were combined and means and standard deviations for the average times on the separate sub-tests were computed.[10] Although all the distributions were similar, approximately the "normal" curve, the ranges of the scores were different and the magnitude of the means and standard deviations varied. In order to combine all the separate averages into a single rate score, it was necessary to take account of these differences lest some tests count more heavily than others in the final measure. Table 11 shows the means and standard deviations[17] of the various distributions as well as the

TABLE 11

MEANS AND STANDARD DEVIATIONS OF AVERAGE SPEED SCORES ON SEPARATE SUB-TESTS TOGETHER WITH THE WEIGHTS USED IN COMBINING THE SEPARATE SCORES INTO A SINGLE MEASURE OF RATE OF SUCCESSFUL WORK

	Boys				Girls			
Sub-test	No. of cases	*M*	*SD*	Weight	No. of cases	*M*	*SD*	Weight
1	93	5.8	1.7	1.00	97	6.6	2.7	1.00
2	92	11.0	3.4	.67	99	11.9	3.6	.75
4	92	8.8	2.4	1.00	95	9.1	2.6	1.00
7	93	7.1	2.1	1.00	98	7.2	2.1	1.50
9	92	6.8	1.8	1.00	98	6.7	1.9	1.50
10	90	15.1	5.6	.33	95	14.4	7.9	.38

---

[10] Two extreme measures, one on Sub-test 1 for boys, the other on Sub-test 4 for girls, were not considered in these calculations.

[17] It will be noted in Table 11 that the numbers of individuals represented in each group again fall short of the total number of children tested. Reasons for omissions of certain items and sub-tests for various subjects have been suggested above but need to be reviewed and supplemented here. For two boys and three girls no final rate score was available because the whole shortened test was in some way invalidated. The two boys were those who took the same form of the test twice, the second one being that given under standard time conditions. For two of the girls the time records were not usable because of failure of the apparatus, and for the other all test scores were discarded because of her reading difficulty.

In certain cases speed scores on one or more sub-tests could not be used. Approximately nine and one-half per cent of the total number of subjects were affected in this way, four and one-half per cent because of irregularities on the first day of testing, and the rest because of inadequacies in the apparatus or in the technique of recording. A total of 37 sub-tests were eliminated from the records of these 19 children.

The number of items on which the final rate score was based was slightly reduced for 29 per cent of the subjects by their failure to indicate omissions. When a child passed an item, after working on it and giving no answer, it was impossible to tell when he had actually begun work on the next problem. Speed measures on a total of 120 items done correctly were weights used in combining individual averages from the separate sub-tests. For the boys, the standard deviations of Sub-tests 1, 4, 7, and 9 are seen to be roughly equivalent while those of 2 and 10 are higher. The weights chosen for these latter two are such that the product of the weight and the standard deviation approximates the sigma of the other four distributions. In the case of the girls, the standard deviations of scores on Sub-tests 1 and 4 were used as the basis and weights were assigned to render the others equivalent. In finding each subject's final score the sum of the weighted averages was divided by the number of sub-tests involved.

Table 12 gives the distribution of the number of items on which the rate scores were actually based. It will be seen immediately that a few of the scores are based on a very small number of items, hence that some of the children are much less reliably measured in respect to speed of successful work than are others. The scores for these subjects were retained since their exclusion would not materially alter the results. Table 13 shows the distributions of the measures

TABLE 12

DISTRIBUTIONS OF THE NUMBER OF ITEMS SERVING AS A BASIS FOR THE RATE OF SUCCESSFUL WORK

No. of items	No. of cases	
	Boys	Girls
70—74	4	
65—69	4	2
60—64	10	4
55—59	8	9
50—54	13	11
45—49	14	16
40—44	15	17
35—39	14	11
30—34	5	6
25—29	3	14
20—24	0	8
15—19	0	3
10—14	2	
5— 9	2	
Total number of cases	94	101

---

thus lost to the records of these 58 children. In a few instances one or more of the items done correctly by a given subject were improperly timed, but only 20 items were eliminated on this account and these were distributed among the records of 11 children representing about four and one-half per cent of the total sample.

TABLE 13

DISTRIBUTION OF SPEED SCORES

Weighted average of number of seconds spent on items done correctly.

Score	Boys	Girls
3.0— 3.9	1	
4.0— 4.9	6	2
5.0— 5.9	23	6
6.0— 6.9	24	14
7.0— 7.9	18	24
8.0— 8.9	16	19
9.0— 9.9	4	14
10.0—10.9	2	5
11.0—11.9		9
12.0—12.9		6
13.0—13.9		1
14.0—14.9		0
15.0—15.9		1
Number	94	101
Mean score	6.6	8.4
SD	1.4	2.1
Coefficient of variation	2.1	2.5

of rate of successful work together with the means, standard deviations and coefficients of variation for the two sexes. The girls tend to be a little slower than the boys and to show greater variability in the speed scores.

3. *Ratings of Behavior in the Test Situation*

Although no use was made of the ratings for that portion of the study presented here, something may be said of the general character of the distributions and of the value of ratings as made under the conditions of this investigation.[18]

For each separate item of behavior, the average rating falls at 4 or 5 on the seven-point scale for both sexes and for both first and second test. Little difference is found between distributions of ratings made on the same trait on the two occasions. This may indicate consistency on the part of the rater, or of the individual subjects, or both. On the other hand it may show only that, in a fairly large group, distributions of ratings for the total sample made at different times show the same general shape. There appear to be

---

[18] See Appendix *A* for rating scale.

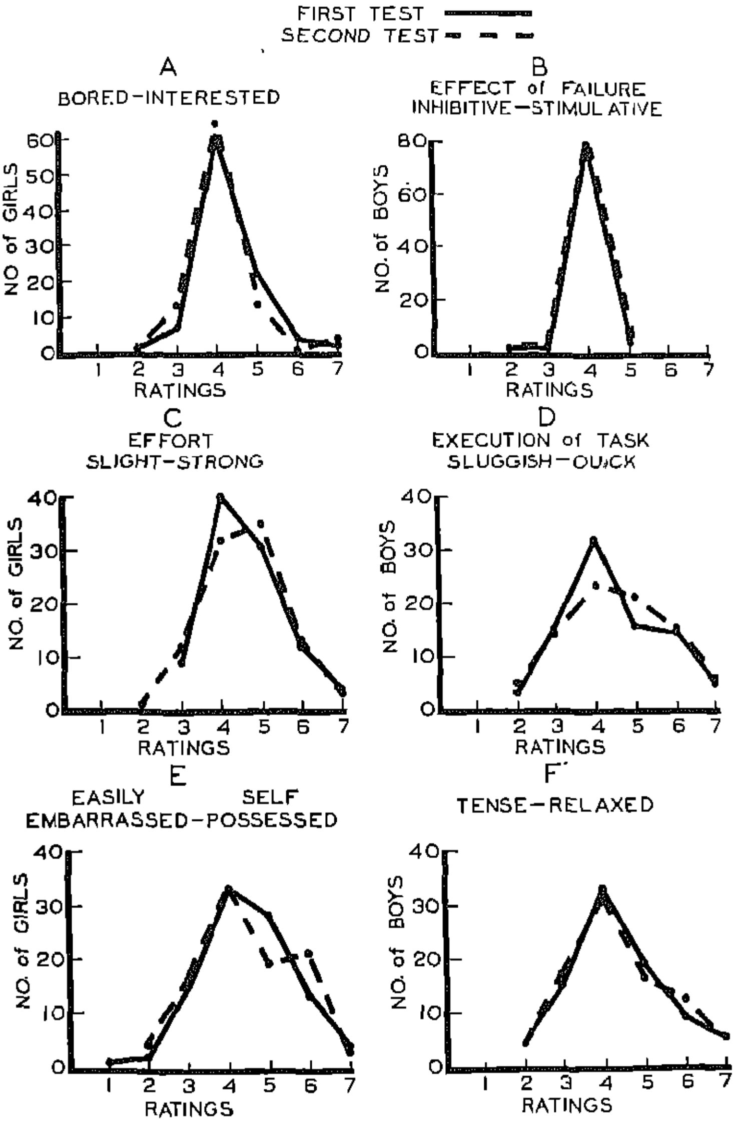

FIGURE 1

Sample Distribution of Ratings Obtained during Administration of the Terman Group Test no outstanding differences in the curves for boys and girls. Rather there is general similarity in range and shape of most of the distributions and close agreement as to those traits which were not adequately rated.

It is to be remembered that the items of the scale were taken from one which had already been used successfully in individual testing situations with preschool children. At least four of the traits—those concerned with interest, eagerness, response to difficulty, and effect of failure—could not be satisfactorily rated under the conditions of the present experiment, the great majority of the ratings falling at the middle or average level of the scale. Graphs *A* and *B* in Figure 1 illustrate this point. The fact that there were no extreme cases at the lower end of the scale and but few at the higher end for both interest and eagerness is probably a reflection of the group's familiarity with and rather nonchalant acceptance of the test situation. The bunching of ratings on response to difficulty and effect of failure at the middle level is undoubtedly due in part to the nature of the test itself. The child, once started on a subtest, proceeded to the end without interruption from the examiner. In the presentation and performance of the tasks there was none of the exchange of conversation which characterizes the more typical individual testing situation. There was no obvious demonstration of inability to respond and little possibility of immediate recognition of all failures either by child or examiner. It seems likely that under these circumstances the subjects were less impressed by the difficulty of the items and less aware of failures than they might have been if another instrument had been used.

In general, use of the group examination reduced the opportunity to observe test behavior since it resulted in the narrowing of the performance to that appropriate for a pencil-and-paper test. Although ratings on effort and facility of performance show a fairly good spread, it was felt that these aspects of behavior were less satisfactorily rated than were those concerned with the child's response to the general situation. Graphs *C, D, E,* and *F* in Figure 1 show distributions of ratings on traits from these sections of the scale.

4. *Short Sample Records*

Another set of data, interesting as descriptive of the group, are those obtained from the short sample records. As has been stated before, these records were taken during only one of the examinations so that no estimate of their consistency is available. Table 14 shows distributions for both boys and girls of the number of 30-second intervals in which various types of behavior were observed during the actual progress of the test. In taking the records, more specific notation was made of the kinds of behavior observed and these were later classified under the headings shown in the table. Face movements, for example, included frowns, grimaces, biting of lips, and the like, while under hand movements were listed such items as picking at the paper, twirling the pencil, and so on. Limb movements included changes in position of leg or arm or shrugging of the shoulder, and body movement covered all shifts of general bodily position. Sex differences are clear only in the case of limb movement, where boys show greater restlessness than do girls, but there is a tendency for the boys to make the top scores in most of the different types of behavior noted.

TABLE 14

DISTRIBUTIONS OF THE TOTAL EXPERIMENTAL GROUP ACCORDING TO THE NUMBER OF 30-SECOND INTERVALS IN WHICH VARIOUS TYPES OF BEHAVIOR WERE NOTED

No. of inter-vals	Type of movement								Yawn, sigh		Vocal.		Appr. to E.	
	Body		Limb		Hand		Face							
	B*	G*	B	G	B	G	B	G	B	G	B	G	B	G
0—4	56	73	11	22		5	58	70	86	92	94	101	80	89
5—9	21	22	13	24	1	8	16	19	9	10		1	11	10
10—14	11	4	17	21	12	13	8	8	1	2	1		4	4
15—19	7	5	13	15	10	15	7	6				2		1
20—24			19	12	22	17	5	1			1			
25—29			14	5	9	9	1							
30—34	1		4	3	14	12	1							
35—39			5	2	9	10							1	
40—44					7	6								
45—49					2	4								
50—54					3	1								
55—59					4	1								
60—64					3	2								
65—69						1								
Number	96	104	96	104	96	104	96	104	96	104	96	104	96	104
Mode	2	2	22	7	22	22	2	2	2	2	2	2	2	2

*B—Boys; G—Girls.

5. *Answers to Questions*

As was mentioned in the section on procedure, certain questions were asked each child at the time of the complete test given under unlimited time conditions. Answers to the question, *"How do you like this kind of a test?"*, were rated on a five-point scale the middle point of which represented an indifferent or non-committal response. Results from the question are presented in Table 15 according to the per cent of children giving each kind of answer. It will be remembered that roughly half the group had the complete test as the first examination while the rest had it as the second. The per cent of cases giving each answer is presented separately for these two subgroups to reveal any differences in the responses of children who had had previous experience with the alternate form of the same test. Fewer children profess to like the test very much when they are asked on the occasion of the second examination and more admit to disliking it, but the distributions are essentially the same and do not differ much from the percentages representing the responses of the total group.

TABLE 15

DISTRIBUTION OF ANSWERS TO QUESTION:
"How do you like this kind of a test?"

Response	Per cent of cases Asked on first test	 Asked on second test	 Total group
Likes very much	10.5	3.5	7.2
Likes	29.5	25.6	27.6
Indifferent	52.6	57.0	54.7
Dislikes	5.3	12.8	8.8
Dislikes very much	2.1	1.2	1.7
Total	100.0	100.1	100.0

TABLE 16

DISTRIBUTION OF ANSWERS TO QUESTION:
"Would you rather take it alone like this or with the whole class?"

Response	Per cent of cases
Alone	43.1
Indifferent	45.7
Group	11.2
Total	100.0

Table 16 shows the per cent of the total group preferring to take tests alone or with the group or expressing indifference in the matter. It is likely that the smallness of the percentages expressing preference for the group situation or admitting dislike for the test is due in part to the subjects' politeness or to their willingness to conform to the conditions presented.

6. *Standard Scores from Intelligence and Achievement Tests*

Standard scores on a variety of intelligence examinations and on the *Stanford Achievement Test* were available for the 196 subjects who were members of the adolescent study group. In calculation of these scores the mean of the distribution for each test was given a value of 50 and individual scores were expressed in terms of their *distance from the mean in standard deviation units, 10 points on* the scale being equal to one sigma. Since these scores are used in some of the calculations to follow, a list of the tests and the dates of administration are presented in Table 17.

TABLE 17

TESTS ADMINISTERED TO EXPERIMENTAL GROUP WITH DATES OF ADMINISTRATION

	Higher grade	Lower grade
CAVD point score (2-hour test)	Spring 1932	Fall 1932
CAVD level score	Spring 1933	Fall 1932
Kuhlmann-Anderson	Fall 1932 Fall 1934	Spring 1933 Spring 1935
Stanford-Binet (short scale)	Spring 1933	Fall 1933
Stanford Achievement (Forms *V* and *W*) Reading and Arithmetic	Fall 1932 Fall 1933 Fall 1934 Fall 1935	Spring 1933 Spring 1934 Spring 1935 Spring 1936
Terman Group Test (Forms *A* and *B*)	Spring 1935	Fall 1935

B. SPEED AND ALTITUDE IN INTELLIGENCE TEST PERFORMANCE

In order to investigate the relationship between speed and altitude as defined and measured in the present experiment, the speed score, representing rate of successful work on one form of the Terman *Group Test,* was correlated with the point score obtained in the other form of the same examination under conditions of unlimited time. It could, perhaps, be argued that this latter measure is not a good indication of altitude since the Terman test is designed for administration under conditions of limited time and might not contain items difficult enough to be an exhaustive measure of level of ability. It was assumed, however, that since the examination is intended for use with students from grades seven to twelve, it would constitute an altitude test for low seventh graders when given in such a manner that each child could attempt every item.

Further measures of level of ability were available and were correlated with rate of successful work. The shortened Stanford-Binet scale was administered in the same semesters as the Terman test, and level scores from the *CAVD* examination had been obtained for most of the group within the year preceding the testing for the present study. The first of these lays little stress on speed of performance, the second is designed for administration without time limits.

Correlations were computed by the product-moment method and are presented in the upper half of Table 18, while related statistical constants are to be found in Table 19. Probable errors ranged from .066 to .069. The relationships revealed are slight, in no case equal

TABLE 18

CORRELATIONS BETWEEN RATE OF SUCCESSFUL WORK AND SCORE ON VARIOUS TESTS OF INTELLIGENCE FOR THE EXPERIMENTAL GROUP

Tests	Score used	Test method	Boys *r*	Girls *r*
Altitude tests				
Termon Group Test (unlimited time)	Point score	Individual	.158	.096
Stanford-Binet (short scale)	Mental age	Individual	.167	.158
CAVD (level)	Standard score*	Group	.224	.147
Speed tests				
Terman Group Test (standard time)	Point score	Individual	.345	.270
Kuhlmann-Anderson	Mental age	Group	.386	.188
CAVD (2-hour test)	Point score	Group	.410	.192

*This was used rather than the level score because of irregularity in dates of testing the higher and lower grades.

TABLE 19
STATISTICAL CONSTANTS FOR ALTITUDE AND SPEED TESTS

Tests	Score		Boys			Girls	
		N	*M*	*SD*	*N*	*M*	*SD*
Altitude tests							
Terman Group Test (unlimited time)	Point	91	101.1	30.9	101	86.7	24.9
Stanford-Binet (short scale)	*MA* in years	90	13.2	1.2	99	12.8	1.3
CAVD (level)	Stand.	90	50.0	9.4	99	49.3	9.8
Speed tests							
Terman Group Test (standard time)	Point	91	90.9	29.1	101	76.6	23.8
Kuhlmann-Anderson	*MA* in years	94	12.3	1.1	99	12.2	.9
CAVD (2-hour test)	Point	90	151.7	26.8	96	144.9	25.7

to four times their probable errors, but the coefficients are consistently positive and uniformly larger for boys. The correlation plots were arranged so that a positive coefficient indicates a positive relationship between high score and fast work. In view of these results it may be stated that the relationship between speed and altitude as determined in the present study is slight but positive.

It will be remembered that the investigation of the speed-altitude relationship grew out of a consideration of the relative merits of individual and group tests, the assumption being that the former were more likely to be tests of altitude, the latter, tests of speed. The findings so far indicate that introduction of the speed factor into the testing situation would not lead to added validity of an intelligence test, but it is yet to be determined whether time limits actually reduce the validity of such an examination.

Scores were available on three "speed" tests of intelligence. The first of these was the Terman *Group Test* administered under conditions of unlimited time but scored according to progress at the end of the standard time allowance. The objection might be made that this score was not representative of performance on a "speed" test since the subjects were not working under a time limit. Attention is called to the revised instructions for the test, described above, and to correlations, to be reported later in another connection, of .848 for boys and .831 for girls,[19] between point scores in standard time on different forms of the same sub-tests administered with and without time limits. These correlations are taken to indicate that children maintain a somewhat similar rank order for scores obtained in a given time whether or not they are working under an announced time limit.

Other "speed" test scores consisted of mental ages from the Kuhlmann-Anderson examination and point scores obtained from the *CAVD* within a 2-hour time limit. The former had been administered six months before the present experiment, the latter, one year before. Correlations of the rate score with these measures and with the point score from the Terman *Group Test* under standard time allowance are presented in the lower half of Table 18. Probable errors for the coefficients vary between .059 and .066.

These correlations show the same trend as was revealed in the comparison between rate and altitude scores but are higher than the latter, all coefficients for boys and one of those for girls being statistically significant. Rate of successful work is thus seen to bear a small but positive and consistent relation to scores on "speed" tests of intelligence. To a certain extent, the faster a child works, the higher will be his score. The relationship is so small, however, as to point to the influence of many other factors beside speed in determining score on such examinations. Therefore, while a child might be penalized to some extent by time limits on this sort of test, his score would probably not be completely misleading as an indication of intelligence.

Further evidence that scores on a so-called speed test result in about the same ranking in respect to intelligence as do scores on an altitude test is furnished by the correlation between score on the complete Terman test within standard time and the unlimited-time, or altitude, score.[20] These coefficients are as follows:

Boys	.958	±	.006
Girls	.940	±	.008

---

[19] These coefficients were obtained by correlating score on seven sub-tests of the first examination with score on seven sub-tests of the second. Points earned within the standard time limits were used, but for each subject one score was obtained when the test was being administered under standard timing conditions, the other while it was being given with unlimited time.

[20] These two scores were obtained at the same sitting.

Granting that the magnitude of the correlation depends in part on the fact that some of the same items are involved in both scores and that these two measures were obtained at the same sitting, it is still true that if a child's score is cut down to that which he obtained within the standard time limit, his rank is about the same as it is when determined by score in unlimited time.

It will be noted that raw rather than corrected coefficients of correlation are presented in the comparisons in Table 18. In view of the high reliabilities of the measures used, this procedure seems justified. The Spearman-Brown reliability of the measure of total time spent on items done correctly on the Terman *Group Test* was over .90 for six sub-tests combined, and it is reasonable to suppose that the rate score, representing a weighted average of these measures, is equally reliable. Thorndike (144, p. 570) estimated the self-corelation of the total *CAVD* score to be near .95 for a group of 144 pupils in a sixth grade. Assuming that the coefficient might be somewhat lower for the sexes separately and for the *CAVD* under restricted time, it is still likely that both the *CAVD* measures here used are highly reliable. As will be reported later, the Spearman-Brown reliabilities for the two Terman test measures are .90 or above. The reliability of the *Stanford Revision* is variously given as from .80 to .95 (53, p. 10). Kuhlmann has not presented reliability coefficients for the Kuhlmann-Anderson examination, but Greene and Jorgensen report it to be an "unusually reliable measurement of mental ability" (61, p. 267).

Terman scores in standard and unlimited time are based on the 10 sub-tests of one form of the complete examination, while rate scores are derived from six of the sub-tests of another form administered at another time. That this does not affect the results significantly is suggested by the correlation between rate of work and point score on the same sub-tests from which the speed measure was derived. These coefficients are closely comparable to the ones found between rate and Terman score for the complete test in standard time and are as follows:

Boys	.341	±	.062
Girls	.317	±	.061

The results reported on the relationship between rate of successful work and scores on intelligence test appear to be independent of test content or method of administration. Speed as measured on items of the Terman *Group Test* is no more closely related to total score on this examination than to score on the other standard tests, the implication being that the rate measure is not exclusive to the Terman test. The general conclusions are the same whether scores obtained through individual or through group procedures are considered.

That rate of successful work is a measure distinct from amount of time needed to complete a test is demonstrated by the correlation obtained between the speed score and number of seconds taken to complete six sub-tests under conditions of unlimited time. The coefficients are as follows:

Boys	.677	±	.039
Girls	.663	±	.038

These correlations indicate that some subjects whose rate of successful work is fast take longer to finish, probably because of time spent on items done incorrectly or on items worked on but not answered, whereas some slow workers take a relatively short time to complete the total test because of hurried or uncritical attack on items not answered.

Before summarizing the findings on speed and level in intelligence test performance it will be pertinent to investigate the reliability and validity of the Terman *Group Test* when given as a timed test and when employed as a measure of altitude. Spearman-Brown reliability coefficients based on odd-even correlations were calculated for the Terman point scores of 58 boys and 68 girls and are presented in Table 20.[21] The scores used are those obtained in standard and in unlimited time when the subjects were working without time restrictions, hence they represent first test scores for about half the group and second test scores for the rest. The coefficients are all of the order of .91 and have probable errors between .021 and .025. There is no evidence here that an examination administered with time limits is less reliable than the same test would be if given as an altitude test.

---

[21] This sub-group of the total sample was matched with another group of children for purposes of a study on social facilitation. Comparison between the total experimental group and this sub-sample revealed no large, significant, or consistent differences between the two.

TABLE 20

RELIABILITY AND VALIDITY OF THE TERMAN GROUP TEST ADMINISTERED WITH AND WITHOUT TIME LIMITS

Coefficient	Terman Group Test Standard time Boys	 Girls	 Unlimited time Boys	 Girls
Spearman-Brown reliability coefficient	.917	.910	.906	.915
Validity coefficient with composite of intelligence test scores as criterion	.803	.799	.802	.807
Validity coefficient with composite of achievement test scores as criterion	.836	.790	.856	.804

The validity coefficients were obtained for the total experimental group, with some reduction in the number of cases due to incomplete records for a few of the subjects. Two criteria for validity were set up, the first being a composite of standard scores on four intelligence tests administered over a period of six school terms. The measures used were the *CAVD* given in limited time, the Stanford-Binet, and two scores from the Kuhlmann-Anderson test, which meant that group and individual, speed and non-speed conditions were all represented. Dates of administration for these tests are included in the section on presentation of the data. The average standard score in the four examinations was correlated with standard score on the complete Terman *Group Test* obtained in limited and in unlimited time.

The second criterion of validity was based on standard scores from four administrations of the *Stanford Achievement Test*. The reading and arithmetic sections of this examination were given every year for four years, and the standard score for each section for a given year represents an average of performance on Forms *V* and *IV*. Yearly standard scores for reading and arithmetic were averaged separately, and the mean of these two final scores was used as the composite achievement measure, being correlated with standard score on the Terman test earned with and without time limits. Spearman-Brown reliability coefficients were available from the Institute records for each year's average score on Forms *V* and *IV* for reading and arithmetic combined. These coefficients, reported for the sexes separately, range from .941 to .972 and are based on scores of the members of the adolescent sample and their classmates, numbers of cases varying from 154 to 165.

The validity coefficients, which have probable errors of from .019 to .025, are presented in Table 20, and statistical constants related to the criterion measures are presented in Table 21. When determined by comparison with other intelligence tests, the validity of the Terman *Group Test* is the same whether time limits are applied or not, all coefficients being close to .80. When determined by comparison with achievement scores, the results are the same, although the coefficients for boys are a little higher than those for girls. Clearly there is nothing to be gained either in reliability or validity by extending the time allowance on this test standardized for use with time limits.

TABLE 21

STATISTICAL CONSTANTS FOR STANDARD SCORE COMPOSITES SERVING AS VALIDITY CRITERIA

	Boys			Girls		
Measure	No. cases	Mean	*SD*	No. cases	Mean	*SD*
Intelligence composite	90	49.5	9.2	102	49.4	8.8
Achievement composite	90	49.4	9.1	102	49.6	8.8

To summarize briefly the findings of this section of the study, it may be said that the relationship between rate of work on an intelligence test and level of ability is insignificant but consistently positive. The relationship between speed of performance and score on a test of intelligence administered with time limits is positive but of small degree. It does not permit the assumption that the score on such an examination is primarily an indication of rate of work rather than level of ability. When the same test is given with and without time limits there is no evidence for greater reliability or validity of scores obtained in unlimited time.

V. SUMMARY AND CONCLUSIONS

Consideration of the general problem of the relative merits of group and individual intelligence testing procedures led to analysis of factors which might influence performance in the two situations. Among the factors discussed, that of rate of work was selected for further study because of conflicting findings from other investigations.

Two hundred boys and girls in the low seventh grade acted as subjects for the present experiment. Both forms of the Terman *Group Test* were administered to each child as individual examinations, one without time restrictions and one with the standard time limits. During the latter test, the time taken on each individual item was recorded for every child so that a measure of rate of successful work was available for all subjects.

One of the criticisms of the group testing procedure has been that the restrictions of time usually necessary when subjects are examined collectively penalizes the slow but able thinker. There are those, however, who claim that the able worker is the fast worker, hence that imposition of time limits should not affect scores on tests of intelligence. Experimental studies designed to attack this problem of the relationship between rate and level of mental ability have led to conflicting results, some showing a fairly high relationship between the two, others showing little or none.

The studies bearing most closely on the present problem may be criticised for several reasons: (*a*) inadequate measures of speed and altitude; (*b*) heterogeneity of the sample; (*c*) narrowness of the function tested. The present study was designed to yield an adequate speed measure based on performance on a standard test of intelligence and an altitude measure derived from the same examination. The speed or rate score was computed on the basis of the average amount of time spent on items done correctly while the altitude measure was the score obtained when all time limits were removed. The rate score was correlated with this and other altitude measures to show the relationship between speed and level of mental ability, and with scores on so-called speed tests of intelligence to indicate the relation of rate of successful work to score on this type of examination. With composites of intelligence and achievement tests as criteria, the relative validity and reliability of the same test given with and without time limits was investigated. Results are as follows:

1. Rate of successful work as measured by average amount of time spent on items done correctly on the Terman *Group Test* bears an insignificant but consistently positive relationship to altitude as measured by this and other tests of intelligence. Correlations range from .096±.069 to .224±.066 and are slightly higher for boys than for girls. The results appear to be independent of test content and method of administration and lead to the conclusion that whatever relationship does exist between speed and altitude is positive.

2. Correlations between rate of successful work and score on various "speed" tests of intelligence, including the Terman examination, range from .188±.066 to .410±.059. Again the coefficients are slightly higher for boys than for girls. Results are independent of test content and method of administration and indicate a small, positive, and for boys significant, relationship between rate of work and score on a "speed" test. The coefficients are too small to indicate that such examinations are purely speed tests. The conclusion is that on the average the influence of time restrictions is fairly negligible.

3. *Neither the validity nor the reliability of the Terman Group Test,* as determined in the present investigation, is affected by administering the examination without time limits. It may be concluded that the validity and reliability of this particular "speed" test of intelligence is not adversely influenced by the fact that it is administered with time restrictions.

The conclusion from the present study which is significant for psychological theory is that which concerns speed as a factor in intelligence. This investigation lends support to the theory that speed and level of mental ability are related; the relationship, although positive, is of a low degree. From the point of view of mental test technology there is little indication that standard group tests place an excessive premium on speed of performance.

APPENDICES

Appendix A

RATING

1. *Coöperativeness:*
    1.1 Attitude toward examiner:
        Negative—friendly ______________________________
        Reticent—garrulous ______________________________
    1.2 Attitude toward situation:
        Bored—interested ______________________________
        Sullen—cheerful ______________________________
    1.3 Attitude toward task:
        Unwilling—eager ______________________________
2. *Effort:*
    2.1 Intensity of effort:
        Slight—strong ______________________________
    2.2 Maintenance of effort:
        Easily discouraged—persistent ______________________________
    2.3 Variability of effort:
        Variable—consistent ______________________________
3. *Facility:*
    3.1 Comprehension of task:
        Slow-swift ______________________________
    3.2 Execution of task:
        Sluggish—quick ______________________________
    3.3 Method of performance:
        Random—systematic ______________________________
4. *Inhibition:*
    4.1 Shyness in situation:
        Timid—unreserved ______________________________
    4.2 Emotional tension:
        Tense-relaxed ______________________________
    4.3 Inertia in specific responses:
        Hesitant—confident ______________________________
    4.4 Response to difficulty:
        Defensive—aggressive ______________________________
    4.5 Effect of failure:
        Inhibitive—stimulative ______________________________
5. *Poise:*
    5.1 Physical poise:
        Fidgety—quiet ______________________________
    5.2 Mental poise:
        Easily embarrassed—self-possessed ______________________________
6. *Attention:*
    6.1 Distractibility:
        Distractible—single minded ______________________________

DEFINITIONS OF POINTS ON RATING SCALE

Scale point	Code number of trait				
	1.1*a*	1.1*b*	1.2*a*	1.2*b*	1.3
1	Completely resistant	Monosyllabic Volunteers nothing	Obviously bored	Definitely sullen	Has to be forced to continue Obvious distaste Open resistance
2	Resists but not completely	Much prodding needed for adequate answer	Dislikes it but endures it	Gloomy	Anxious to be through. Watches time. Won't go on. Counts tests
3	A little resistant	Prodding needed for full answer	Not interested Waste of time	No trace of cheerfulness Solemn	Not willing but submitting as gracefully as possible
4	Neither one nor other Natural	Talks fully as far as called for	Thinks it is "all right"	Neither one nor other	Nothing positive but perfectly willing
5	More friendly than necessary	Volunteers. Takes advantage of openings	Takes advantage of opportunity to establish social relations	Definitely cheerful	Rather enjoying it
6	Some reserve Not quite as open at 7	Talks much but does not have to be stopped	Makes opportunity to establish social relations	Smiling and laughing a good deal	Definitely enjoying it
7	Extremely responsive	Has to be quelled	Has a good time	Obviously in high spirits	Anxious for more

DEFINITIONS OF POINTS ON RATING SCALE

Scale point	Code number of trait					
	2.1	2.2	2.3	3.1	3.2	3.3
1	Tries not at all. Takes path of least resistance	Can't be encouraged to try	Effort varies from item to item in same sort of test	Has to be told over and over	All tests much over time	Heedless Slap-dash
2	Makes slight effort if it seems easy or enjoyable	Has to be made to stay with problem	Effort varies frequently from test to test	Says he understands but does not. Asks many questions	All tests overtime	
3	Puts forth minimum of effort	Has to be encouraged to give due consideration	Noticeable change of effort from test to test	Does not fully understand or asks many questions	0-1 test finished early	
4	Makes adequate attempt to do as requested	Gives due consideration Willing to give up if too hard	Spurts and slumps occasionally	Asks a question or two, or for repetition of directions	2-3 tests finished early	
5	Makes better than satisfactory effort	Unwilling to give up without work on it	Definitely consistent	No questions necessary. Not particularly quick	3-5 tests finished early	
6	Gives evidence of tension and concentration	Persists until absolutely convinced can't do it	Shows slight variations for obvious reasons	Understands immediately Does not anticipate	5-7 tests finished early	
7	Shows marked tension and concentration	Persists in spite of obvious inability to do it	Attacks all problems with same effort	Goes right ahead. Does not wait for directions	8-10 tests finished early	Goes very carefully Erases errors

DEFINITIONS OF POINTS ON RATING SCALE

Scale point	Code number of trait 4.1	4.2	4.3	4.4	4.5
1	Afraid to act or talk	Nervous laughter Trembling hands or voice. Working under load. Uneasy	Afraid to utter positive statement. Makes defensive remarks	Every response prefaced or followed by excuse	Needs much encouragement
2		Shows some of above symptoms	Vacillates Wants approval Gives tentative answers	Makes excuses frequently	Needs some encouragement
3		Continued slight strain throughout	Some hesitation Some withholding of response	Makes excuses once or twice in whole test	Slowed up and subdued after failure
4	Not timid but restrained	Slight but not continuous strain	Unhesitatingly submits answer for what it is worth	Takes test in stride	Not affected by failure
5	Not restrained Does not take initiative	Occasional strain	Feels answer is probably right	Slight spurt when difficulty appears	Tries harder after failure
6	Shows momentary reserve	Rare strain	Feels sure of self	Vigor of attack apt to increase with difficulty	Apt to do best work after failure
7	No hesitation about talking or acting naturally	Calm. At ease Quiet	Expresses confidence in his own ability	Vigor of attack increases with difficulty	Consistently does best work after failure

DEFINITIONS OF POINTS ON RATING SCALE

Scale point	Code number of trait		
	5.1	5.2	6.1
1	Walks about. Kneels on chair. Squirms Twists	Blushes. Laughs nervously Looks sheepish. Talks little or much. Teary. Turns away Self-conscious	Distracted by slight noises and for no apparent reason
2	Rarely quiet	Disturbed continuously Shows some of above signs	Responds to every noticeable stimulus
3	Continuous vigorous or varied movements	Disturbed often but carries on	Responds to nearly all external stimuli
4	Large and small movements from time to time	Disturbed sometimes but not permanently. Easily over it	Responds to some outside stimuli Ignores others
5	Some large or small movements. Gives impression of quiet	Some slight disturbance during test	Responds to very loud noises, etc.
6	Occasional small movements. Almost no large movements	Rarely disturbed	No distractions except when there is very disturbing stimulus
7	No unnecessary muscular movement	Completely lacking in self-consciousness	Absolutely oblivious to external stimuli

Appendix B

Revised Instructions for the Terman Group Test

1. *Shortened Test, Standard Time*

Here is a test to see what pupils can do. Don't open it until I tell you to. Now fill in your name on the top line of the first page.

Listen carefully to what you are told to do. Don't begin until I say *"Go."* When I say *"Stop"* you must stop instantly. Go as fast as you can without hurrying. Try each question in the right order, but don't spend too much time on one you don't understand. If it's really too hard to do, make a circle around the number and go right on to the next one.

Now turn over the page to Test 1 and fold your book back this way.

(Directions for the separate subtests were not changed except for insertion of the sentence *"Remember to draw a circle around the numbers of the ones that are really too hard to do"* before the signal to begin.)

2. *Total Test, Unlimited Time*

(Same as above except for omission of third sentence of second paragraph. Instructions to circle numbers were not given before each separate subtest and the subjects were not stopped at the end of the standard time allowance.)

3. *Group Test for Matched Control Sample*

Here is a test to see what pupils can do. I will give each of you a test book. Do not open it until I tell you to.

Now fill in your name on the top line of the first page.

Attention. Listen carefully to what you are told to do. Ask no questions. Do not look at your neighbor to see what he does. Do not begin until I say *"Go."* When I say *"Stop"* you must stop instantly and hold your pencils up. Go as fast as you can without hurrying. Try each question in the right order, but don't spend too much time on one you do not understand. If it's really too hard to do, make a circle around the number and go right on to the next one. (Rest of directions as in manual.)

REFERENCES

1. Abelson, A. R. The measurement of mental ability of "backward" children. *Brit. J. Psychol.*, 1911, **4**, 268-314.
2. Aikens, H. A., Thorndike, E. L., & Hubbell, E. Correlations among perceptive and associative processes. *Psychol. Rev.*, 1902, **9**, 374-382.
3. Allport, F. H. The influence of the group upon association and thought. *J. Exper. Psychol.*, 1920, **3**, 159-182.
4. ———. Social Psychology. Cambridge: Riverside Press, 1924. Pp. xiv+453.
5. Allport, G. W. Personality: A Psychological Interpretation. New York: Holt, 1937. Pp. xiv+588.
6. Allport, G. W., & Vernon, P. E. Studies in Expressive Movement. New York: Macmillan, 1933. Pp. xiii+269.
7. Armstrong, C. P. Sex differences in the mental functioning of school children. *J. Appl. Psychol.*, 1932, **16**, 559-571.
8. Atkinson, W. R. The relation of intelligence and of mechanical speeds to the various stages of learning. *J. Exper. Psychol.*, 1929, **12**, 89-112.
9. Attenborough, J., & Farber, M. The relation between intelligence, mechanical ability, and manual dexterity in special school children. *Brit. J. Educ. Psychol.*, 1934, **4**, 140-161.
10. Bagley, W. C. On the correlation of mental and motor ability in school children. *Amer. J. Psychol.*, 1901, **12**, 193-205.
11. Beck, L. F. Relation of speed of reaction to intelligence. *Amer. J. Psychol.*, 1932, **44**, 793-795.
12. Bernstein, E. Quickness and intelligence. *Brit. J. Psychol. Monog. Suppl.*, 1924, No. 7. Pp. 55.
13. Blackburn, J. M. Methods of estimating intelligence and personality and their application. *J. Ment. Sci.*, 1938, **84**, 1008-1053.
14. Bolton, T. L. The relation of motor power to intelligence. *Amer. J. Psychol.*, 1903, **14**, 615-631.
15. Bowen, J. L. Appraising the individual's ability. *J. High. Educ.*, 1933, **4**, 310-318.
16. Bronner, A. F. Variations in mental equipment not indicated by an intelligence ratio. (Abstract). *Psychol. Bull.*, 1920, **17**, 65.
17. Bronner, A. F., Healy, W., Lowe, G. M., & Shimberg, M. E. A Manual of Individual Mental Tests and Testing. Boston: Little, Brown, 1928. Pp. x+287.
18. Brown, R. R. Effect of age on the speed-power relationship with reference to tests of intelligence. *J. Educ. Psychol.*, 1938, **29**, 413-418.
19. Brown, W. Some experimental results in the correlation of mental abilities. *Brit. J. Psychol.*, 1910, **3**, 296-322.
20. Buros, O. K. (Ed.) The 1938 Mental Measurements Yearbook. New Brunswick: Rutgers Univ. Press, 1938. Pp. xv+415.
21. Burt, C. Experimental tests of general intelligence. *Brit. J. Psychol.*, 1910, **3**, 94-117.
22. ———. Experimental tests of higher mental processes and their relation to general intelligence. *J. Exper. Ped.*, 1911, **1**, 93-112.

23. ———. Mental and Scholastic Tests. London: King, 1922. Pp. xv+432.

24. Carlson, W. S. Further studies in expressive movement. *Psychol. Rec.*, 1938, **2**, 310-316.

25. Cattell, J. McK., & Farrand, L. Physical and mental measurements of the students of Columbia University. *Psychol. Rev.*, 1896, **3**, 618-648.

26. Chapman, J. C. Persistence, success and speed in a mental task. *Ped. Sem.*, 1924, **31**, 276-284.

27. Clark, J. R. The Relation of Speed, Range and Level to Scores on Intelligence Tests. Ph.D. thesis, Dept. of Phil., Columbia Univ., 1925 (privately printed).

28. Commins, W. D. Principles of Educational Psychology. New York: Ronald Press, 1937. Pp. xiv+596.

29. Copeland, H. A. Age differences in mental ability as measured by a work-limit test. (Abstract.) *Psychol. Bull.*, 1938, **35**, 642-643.

30. Dashiell, J. F. Experimental studies of the influence of social situations on the behavior of individual human adults. In *Handbook of Social Psychology*. Worcester: Clark Univ. Press, 1935. (Pp. 1097-1158.)

31. Dearborn, G. Van N. The determination of intellectual regression and progression. *Amer. J. Psychiat.*, 1927, **6**, 725-741.

32. Deputy, E. C. Predicting first-grade reading achievement. *Teach. Coll. Contrib. Educ.*, 1930, No. 426. Pp. 61.

33. Dickson, V. E. Mental Tests and the Classroom Teacher. Yonkers-on-Hudson: World Book, 1924. Pp. xv+231.

34. Douglas, O. B., & Holland, B. F. Fundamentals of Educational Psychology. New York: Macmillan, 1938. Pp. xv+598.

35. Dowd, C. E. A study of the consistency of rate of work. *Arch. of Psychol.*, 1926, **13**, No. 84. Pp. 33.

36. Downey, J. E. The Will-Temperament and Its Testing. Yonkers-on-Hudson: World Book, 1923. Pp. v+339.

37. DuBois, P. H. A speed factor in mental tests. *Arch. of Psychol.*, 1932, **22**, No. 141. Pp. 38.

38. Durrell, D. D. The influence of reading ability on intelligence measures. *J. Educ. Psychol.*, 1933, **24**, 412-416.

39. Farnsworth, P. R., Seashore, R. H., & Tinker, M. A. Speed in simple and serial action as related to performance in certain "intelligence" tests. *J. Genet. Psychol.*, 1927, **34**, 537-551.

40. Farson, M. R. A report on the examination of one hundred *6B* children in Philadelphia schools. *Psychol. Clin.*, 1928, **17**, 128-152.

41. Fenton, N., & Worcester, D. An Introduction to Educational Measurements. New York: Ginn, 1928. Pp. ix+149.

42. Foley, J. P., Jr. An experimental study of the effect of occupational experience upon motor speed and preferential tempo. *Arch. of Psychol.*, 1937, **31**, No. 219. Pp. 40.

43. ———. Factors conditioning motor speed and tempo. *Psychol. Bull.*, 1937, **34**, 351-397.

44. Franzen, R. Attempts at test validation. *J. Educ. Res.*, 1922, **6**, 145-158.

45. Freeman, F. N. Note on the relation between speed and accuracy or quality of work. *J. Educ. Res.*, 1923, **7**, 87-89.

46. ———. Mental Tests. New York: Houghton Mifflin, 1926. Pp. ix+503.

47. ———. Mental Tests. (2nd ed.) Boston: Houghton-Mifflin, 1939. Pp. x+460.

48. Freeman, F. S. Power and speed: their influence upon intelligence test scores. *J. Appl. Psychol.*, 1928, **12**, 631-635.

49. ———. The factors of speed and power in tests of intelligence. *J. Exper. Psychol.*, 1931, **14**, 83-90.

50. ———. The factor of speed. *J. Gen. Psychol.*, 1932, **6**, 462-468.

51. Frischeisen-Köhler, I. The personal tempo and its inheritance. *Char. & Personal.*, 1933, **1**, 301-313.

52. Garrett, H. E. Statistics in Psychology and Education. New York: Longmans, Green, 1931. Pp. xiii+317.

53. Garrett, H. E., & Schneck, M. R. Psychological Tests, Methods, and Results, Part 2. New York: Harpers, 1933. Pp. 235.

54. Garrison, K. C. An investigation of some simple speed activities. *J. Appl. Psychol.*, 1929, **13**, 167-172.

55. Gates, A. I. An experimental and statistical study of reading and reading tests. *J. Educ. Psychol.*, 1921, **12**, 303-314; 378-391.

56. ———. The correlations of achievement in school subjects with intelligence tests and other variables. *J. Educ. Psychol.*, 1922, **13**, 129-139; 223-235; 277-285.

57. ———. The unreliability of *MA* and *IQ* based on group tests of general mental ability. *J. Appl. Psychol.*, 1923, **7**, 93-100.

58. Geyer, D. L. The reliability of rankings by group intelligence tests. *J. Educ. Psychol.*, 1922, **13**, 43-49.

59. Gilbert, J. A. Researches on the mental and physical development of school children. *Stud. Yale Psychol. Lab.*, 1894, **2**, 40-100.

60. Greene, H. A., & Jorgensen, A. N. The Use and Interpretation of Educational Tests. New York: Longmans, Green, 1930. Pp. xxii+389.

61. ———. The Use and Interpretation of High School Tests. New York: Longmans, Green, 1937. Pp. xxvi+614.

62. Hargreaves, H. L. The "faculty" of imagination. *Brit. J. Psychol. Monog. Suppl.*, 1927, No. 10. Pp. 74.

63. Hart, B., & Spearman, C. Mental tests of dementia. *J. Abn. Psychol.*, 1914, **9**, 217-264.

64. Healy, W., & Fernald, G. M. Tests for practical mental classification. *Psychol. Monog.*, 1911, **13**, No. 2. Pp. viii+54.

65. Highsmith, J. A. Relation of the rate of response to intelligence. *Psychol. Monog.*, 1925, **34**, No. 3. Pp. 33.

66. Hilden, A. H., & Skeels, H. M. A comparison of the Stanford-Binet scale, the Kuhlmann-Anderson group test, the Arthur point scale of performance tests and the Unit Scales of attainment. *J. Exper. Educ.*, 1935, **4**, 214-230.

67. Hildreth, G. H. Psychological Service for School Problems. Yonkers-on-Hudson: World Book, 1930. Pp. xiii+317.

68. HILL, H. S. Correlation between *IQ's* of bilinguals at different ages on different intelligence tests. *Sch. & Soc.*, 1936, **44**, 89-90.

69. HOLLINGWORTH, L. S., & MONAHAN, J. E. Tapping rate of children who test above 135 *IQ* (Stanford-Binet). *J. Educ. Psychol.*, 1926, **17**, 505-518.

70. HUNSICKER, L. M. A study of the relationship between rate and ability. *Teach. Coll. Contrib. Educ.*, 1925, No. 185. Pp. vi+52.

71. INSKEEP, A. D. Child Adjustment in Relation to Growth and Development. New York: Appleton, 1930. Pp. xiv+427.

72. JUDD, C. H. Measuring the Work of the Public Schools. Cleveland: Survey Committee of Cleveland Found., 1916. Pp. 290.

73. KELLER, H., & WEBER, C. O. Germany's elimination tests for dentists. *J. Appl. Psychol.*, 1932, **16**, 465-475.

74. KELLEY, T. L. Crossroads in the Mind of Man. Stanford University: Stanford Univ. Press, 1928. Pp. vii+238.

75. KENNEDY, M. Speed as a personality trait. *J. Soc. Psychol.*, 1930, **1**, 286-299.

76. KENT, G. H. Sources of error in the use of statistical methods for individual study. (Abstract.) *Psychol. Bull.*, 1924, **22**, 117.

77. ———. Written tests for the clinic. *J. Genet. Psychol.*, 1934, **44**, 49-68.

78. ———. The use and abuse of mental tests in clinical diagnosis. *Psychol. Rec.*, 1938, **2**, 391-400.

79. KENT, G. H., & SHAKOW, D. Group tests for clinical studies. *J. Genet. Psychol.*, 1928, **35**, 595-618.

80. KIRKPATRICK, E. A. Individual tests of school children. *Psychol. Rev.*, 1900, **7**, 274-280.

81. KLINEBERG, O. An experimental study of speed and other factors in "racial" differences. *Arch. of Psychol.*, 1928, **15**, No. 93. Pp. 111.

82. ———. Mental tests. *Encyclopaedia of the Social Sciences*, 1933, **10**, 323-328.

83. LAMBETH, M., & LANIER, L. H. Race differences in speed of reaction. *J. Genet. Psychol.*, 1933, **42**, 255-297.

84. LEMMON, V. W. The relation of reaction time to measures of intelligence, memory, and learning. *Arch. of Psychol.*, 1927, **15**, No. 94. Pp. 38.

85. LINE, W., & KAPLAN, E. The existence, measurement and significance of a speed factor in the abilities of public school children. *J. Exper. Educ.*, 1932, **1**, 1-8.

86. LORGE, I. The influence of the test upon the nature of mental decline as a function of age. *J. Educ. Psychol.*, 1936, **27**, 100-110.

87. LOUTTIT, C. M. Clinical Psychology. New York: Harpers, 1936. Pp. xx+695.

88. MADSEN, I. N. Educational Measurement in the Elementary Grades. Yonkers-on-Hudson: World Book, 1930. Pp. x+294.

89. MAY, M. A. Effect of doubling the time limits in the Alpha and Beta examinations. *Nat. Acad. Sci. Mem.*, 1921, **15**, 415-420.

90. MCCALL, W. A. Measurement. New York: Macmillan, 1939. Pp. xv+535.

91. McFarland, R. A. An experimental study of the relationship between speed and mental ability. *J. Gen. Psychol.*, 1930, **3**, 67-97.

92. Mead, C. The relations of general intelligence to certain mental and physical traits. *Teach. Coll. Contrib. Educ.*, 1916, No. 76. Pp. 117.

93. Meier, N. C. A study of the Downey Test made by the method of estimates. *J. Educ. Psychol.*, 1923, **14**, 385-395.

94. Miles, C. C. Influence of speed and age on intelligence scores of adults. *J. Gen. Psychol.*, 1934, **10**, 208-210.

95. Miles, C. C., & Miles, W. R. The correlation of intelligence scores and chronological age from early to late maturity. *Amer. J. Psychol.*, 1932, **44**, 44-78.

96. Miles, W. R. Abilities of older men. *Person. J.*, 1933, **11**, 352-357.

97. Miller, E. Emotional factors in intellectual retardation. *J. Ment. Sci.*, 1933, **79**, 614-625.

98. Miller, W. S. Variation of *IQ*'s obtained from group tests. *J. Educ. Psychol.*, 1933, **24**, 468-474.

99. Moore, H. A comparison of linguistic and non-linguistic ability. *J. Soc. Psychol.*, 1931, **2**, 245-252.

100. Moss, F. A. Applications of Psychology. New York: Houghton Mifflin, 1929. Pp. x+477.

101. Nemzek, C. The constancy of the *IQ*. *Psychol. Bull.*, 1933, **30**, 143-168.

102. Oates, D. W. A statistical and psychological investigation of intelligence tests. *Forum Educ.*, 1928, **6**, 38-62.

103. Odell, C. W. Educational Measurement in High School. New York: *Century*, 1930. Pp. xiv+641.

104. Orleans, J. S. Measurement in Education. New York: Nelson, 1937. Pp. xvi+461.

105. Page, M. L. The mechanical ability of subnormal boys. *J. Appl. Psychol.*, 1933, **17**, 164-181.

106. Peak, H., & Boring, E. G. The factor of speed in intelligence. *J. Exper. Psychol.*, 1926, **9**, 71-94.

107. Perrin, F. A. C. An experimental study of motor ability. *J. Exper. Psychol.*, 1921, **4**, 24-56.

108. Peterson, J. The comparative abilities of white and negro children. *Comp. Psychol. Monog.*, 1923, **1**, No. 5. Pp. 141.

109. Peterson, J., & Lanier, L. H. Studies in the comparative abilities of whites and negroes. *Ment. Meas. Monog.*, 1929, No. 5. Pp. 156.

110. Peterson, J., Lanier, L. H., & Walker, H. M. Comparisons of white and negro children in certain ingenuity and speed tests. *J. Comp. Psychol.*, 1925, **5**, 271-283.

111. Philip, B. R. Reaction-times of children. *Amer. J. Psychol.*, 1934, **46**, 379-396.

112. Pintner, R. Intelligence Testing. (New ed.) New York: Holt, 1931. Pp. xii+555.

113. Piotrowski, Z. Work habits as a means of differentiating native and acquired mental deficiency. (Abstract.) *Psychol. Bull.*, 1933, **30**, 728-729.

114. Pressey, L. W. The primary classification test. *J. Educ. Res.*, 1924, **9**, 305-314.

115. Pressey, S. L. Distinctive features in psychological test measurements made upon dementia praecox and chronic alcoholic patients. *J. Abn. Psychol.*, 1917, **12**, 130-139.

116. Pritchard, M. C. The mechanical ability of subnormal boys. *Teach Coll. Contrib. Educ.*, 1937, No. 699. Pp. x+73.

117. Rhodes, A. A comparative study of motor abilities of negroes and whites. *Child Devel.*, 1937, **8**, 369-371.

118. Root, W. T. A socio-psychological study of fifty-three supernormal children. *Psychol. Monog.*, 1921, **29**, No. 4. Pp. 134.

119. Rounds, G. H. Is the latent time in the Achilles tendon reflex a criterion of speed in mental reactions? *Arch. of Psychol.*, 1928, **15**, No. 95. Pp. 91.

120. Ruch, G. M. The speed factor in mental measurements. *J. Educ. Res.*, 1924, **9**, 39-45.

121. Ruch, G. M., & Koerth, W. "Power" vs. "speed" in Army Alpha. *J. Educ. Psychol.*, 1923, **14**, 193-208.

122. Ruch, G. M., & Stoddard, G. D. Tests and Measurements in High School Instruction. Yonkers-on-Hudson: World Book, 1927. Pp. xix+381.

123. Saum, A. L. The relative significance of time as a factor in tests of ability at the college level. *J. Appl. Psychol.*, 1938, **22**, 192-210.

124. Schwesinger, G. C. Heredity and Environment. New York: Macmillan, 1933. Pp. viii+484.

125. Seashore, R. H. Individual differences in motor skills. *J. Gen. Psychol.*, 1930, **3**, 38-66.

126. Shen, E. The standard error of certain estimated coefficients of correlation. *J. Educ. Psychol.*, 1924, **15**, 462-465.

127. Simpson, B. R. Correlation of mental abilities. *Teach. Coll. Contrib. Educ.*, 1912, No. 53. Pp. 122.

128. Sisk, T. K. The interrelations of speed in simple and complex processes. *George Peabody Coll. for Teach. Contrib. Educ.*, 1926, No. 23. Pp. 48.

129. Skaggs, E. B. An Elementary Textbook of Mental Measurements. Ann Arbor: Wahr, 1923. Pp. 169.

130. Slater, P. Speed of work in intelligence tests. *Brit. J. Psychol.*, 1938, **29**, 55-68.

131. Spearman, C. E. The Nature of "Intelligence" and the Principles of Cognition. New York: Macmillan, 1923. Pp. viii+358.

132. ———. The Abilities of Man. New York: Macmillan, 1927. Pp. vi+415+xxxiii.

133. ———. *"G" and after—a school to end schools.* In *Psychologies of 1930.* Worcester: Clark Univ. Press, 1930. (Pp. 339-366.)

134. ———. Psychology down the Ages. Vol. 2. London: Macmillan, 1937. Pp. vii+355.

135. Spence, R. B. A comprehensive testing program for elementary schools. *Teach. Coll. Rec.*, 1933, **34**, 279-284.

136. Squire, C. R. Graded mental tests. *J. Educ. Psychol.*, 1912, **3**, 363-380; 430-443; 493-506.

137. Sutherland, J. D. The speed factor in intelligent reactions. *Brit. J. Psychol.*, 1934, **24**, 276-295.

138. SYMONDS, P. M. Mental Hygiene of the School Child. New York: Macmillan, 1934. Pp. xi+321.

139. TERMAN, L. M. The Measurement of Intelligence. New York: Houghton Mifflin, 1916. Pp. xviii+362.

140. ———. The Intelligence of School Children. New York: Houghton Mifflin, 1919. Pp. xxii+317.

141. ———. Terman Group Test of Mental Ability. Manual of Directions. Yonkers-on-Hudson: World Book, 1920. Pp. 12.

142. TERMAN, L. M., *et al.* Mental and Physical Traits of a Thousand Gifted Children. Genetic Studies of Genius, Vol. 1. Stanford University: Stanford Univ. Press, 1926. Pp. xiii+648.

143. TERMAN, L. M., & MERRILL, M. A. Measuring Intelligence. New York: Houghton Mifflin, 1937. Pp. xi+461.

144. THORNDIKE, E. L., *et al.* The Measurement of Intelligence. New York: Teach. Coll., Columbia Univ., 1926. Pp. xxvi+616.

145. THURSTONE, L. L. Ability, motivation and speed. *Psychometrika,* 1937, **2**, 249-254.

146. ———. Primary mental abilities. *Psychomet. Monog.,* No. 1. Chicago: Univ. Chicago Press, 1938. Pp. 121.

147. TINKER, M. A. Speed in intelligence. *J. Appl. Psychol.,* 1934, **18**, 189-196.

148. TRAVIS, L. E., & DORSEY, J. M. The relationship of "intelligence" and reflex conduction rate as found in hypophrenic children. *J. Exper. Psychol.,* 1930, **13**, 370-372.

149. TRAVIS, L. E., & HUNTER, T. A. The relation between "intelligence" and reflex conduction rate. *J. Exper. Psychol.,* 1928, **11**, 342-354.

150. TRAVIS, L. E., & YOUNG, C. W. The relation of electromyographically measured reflex time in the patellar and Achilles reflexes to certain physical measurements and to intelligence. *J. Gen. Psychol.,* 1930, **3**, 374-400.

151. TROW, W. C. Trait consistency and speed of decision. *Sch. & Soc.,* 1925, **21**, 538-542.

152. ———. Introduction to Educational Psychology. Cambridge: Riverside Press, 1937. Pp. x+416.

153. TRYON, C. McC. On the nature of "speed" and its relation to other variables. *J. Gen. Psychol.,* 1933, **8**, 198-216.

154. TRYON, C. McC., & JONES, H. E. The relationship between "speed" and "altitude." *J. Exper. Psychol.,* 1933, **16**, 98-114.

155. URHBROCK, R. S. An analysis of the Downey Will-Temperament tests. *Teach. Coll. Contrib. Educ.,* 1928, No. 296. Pp. 78.

156. VITELES, M. S. The clinical viewpoint in vocational selection. (Abstract.) *Psychol. Bull.,* 1925, **22**, 113-114.

157. ———. The influence of training on motor test performance. *J. Exper. Psychol.,* 1933, **16**, 556-564.

158. WALLIN, J. E. W. A comparison of three methods for making the initial selection of presumptive mental defectives. *Sch. & Soc.,* 1921, **13**, 31-45.

159. ———. Clinical and Abnormal Psychology. New York: Houghton Mifflin, 1927. Pp. xxii+649.

160. WALTERS, F. C. A statistical study of certain aspects of the time factor in intelligence. *Teach. Coll. Contrib. Educ.*, 1927, No. 248. Pp. viii+82.

161. WEBB, E. Character and intelligence. *Brit. J. Psychol. Monog. Suppl.*, 1915, No. 3. Pp. 99.

162. WELLS, F. L. Mental Tests in Clinical Practice. New York: World Book, 1927. Pp. x+315.

163. WHIPPLE, G. M. Reaction-times as a test of mental ability. *Amer. J. Psychol.*, 1904, **15**, 489-498.

164. ———. Manual of Mental and Physical Tests, Part 1. Baltimore: Warwick & York, 1921. Pp. xvi+365.

165. WHITEHORN, J. C., LUNDHOLM, H., & GARDNER, G. E. Concerning the alleged correlation of intelligence with knee jerk reflex time. *J. Exper. Psychol.*, 1930, **13**, 293-295.

166. WIRES, E. M. Interpretation of reactions to the Pintner-Paterson performance scale. *J. Educ. Res.*, 1931, **24**, 53-57.

167. WISSLER, C. M. The correlation of mental and physical tests. *Psychol. Rev. Monog. Suppl.*, 1901, **3**, No. 6. Pp. 62.

168. WOOLLEY, H. T., & FISCHER, C. R. Mental and physical measurements of working children. *Psychol. Monog.*, 1914, **18**, No. 1. Pp. 247.

169. WYATT, S. The quantitative investigation of higher mental processes. *Brit. J. Psychol.*, 1913, **6**, 109-133.

170. YOAKUM, C. S., & YERKES, R. M. Army Mental Tests. New York: Holt, 1920. Pp. xiii+303.

171. ZACHRY, C. B., & LLOYD, W. Interpretation of "intelligence" tests. *J. Nat. Educ. Assoc.*, 1932, **21**, 149-150.

Published as a separate and in *Genetic Psychology Monographs*, 1941, **23**, 319-385.

INDIVIDUAL DIFFERENCES IN APPERCEPTIVE REACTION: A STUDY OF THE RESPONSE OF PRESCHOOL CHILDREN TO PICTURES*

ELISABETH W. AMEN

Institute of Child Welfare, University of California

Acknowledgments 321

I. Introduction 323

II. Experimental problem 325
  A. Statement of problem 325
  B. Experimental materials 325
  C. Subjects 332
  D. Method 334
  E. Illustrative records 337

III. Analysis of data 345
  A. Categories of perceptual response 345
  B. Misinterpretation of details 347
  C. Part-whole balance in perception 352
  D. Fiction play 355
  E. Color interest 364
  F. Projection 365
  G. Summary 365

IV. Individual protocols 367

V. Preliminary evaluation of the procedure 381

VI. Conclusions 383

References 385

---

*Accepted for publication by Harold E. Jones of the Editorial Board, and received in the Editorial Office on February 3, 1941. Copyright by The Journal Press.

ACKNOWLEDGMENTS

The writer wishes to make grateful acknowledgment of the generous assistance of many members of the Institute of Child Welfare Nursery School Staff, University of California, and of the *W. P. A.* Nursery Education Staff of Berkeley, California; and is especially indebted to Dr. Harold E. Jones, Director of the Institute of Child Welfare, for his encouragement of the project, for editorial assistance, and for the services of a student-artist, Miss Betty Byrnes, who made possible the production of the pictures.

Both clerical and photographic assistance in the preparation of this manuscript was furnished by the personnel of Works Progress Official Project No. 665-08-3-30.

ELISABETH W. AMEN

Wheaton College
Norton, Massachusetts

I. INTRODUCTION

The reaction of children to pictures has long been utilized in intelligence tests as an index of apperceptive maturity, regarded from the point of view of the level of intellectual function represented. More recently, another use of pictures has been developed in connection with the study of personality differences revealed in apperceptive responses, as in the Murray *Thematic Apperception Test* (8).

The present study involves an analysis of the responses of 77 preschool children to each of a series of 15 pictures. The pictures were prepared not as items for the measurement of mental ability, but with the purpose of providing material relevant to an "individual's private world of meanings, significances, patterns, and feelings" (Frank, 3, p. 402). This "projective" function was sought by selecting topics approximating the life experiences and interests of preschool children; situations were included which presented possible interpretations in terms of conflict, jealousy, fear, aggression, etc. Consideration was also given to the criteria proposed by Symonds (10) for the selection of pictures in projection studies; these include the requirements that the picture shall present a minimum of detail, shall be vague in theme and incomplete in content, and shall present characters with whom the experimental subjects can readily identify themselves.

II. EXPERIMENTAL PROBLEM

A. Statement of Problem

Although the writer's purpose was to develop a technique for use in projective studies of young children, the present report deals only with preliminary results considered from a normative point of view, and without presenting evidence as to the clinical validity of the technique. Writers in this general field, with high hopes for the diagnostic effectiveness of a specific projective method, have sometimes failed to report results from a normal sampling against which clinical data could be evaluated. This study, then, is concerned with the first steps of (*a*) developing a procedure and (*b*) presenting a normative background for subsequent work.

B. Experimental Materials

Figure 1 shows a photograph of the first picture, a colored background painting of a large table set for a birthday party. At one side of the table are six movable cut-out figures, and on the table are two movable gift boxes. The size of the picture is 11 by 14 inches; the children's figures are five inches in height, and the adults' figures 8½ and 7¾ inches in height. In presenting this picture to the child, the examiner remarks, *"Let's pretend it's someone's birthday . . . whose birthday is it?"* After one of the figures has been chosen as having the birthday, the question is asked, *"Who will come to the party?"* This is followed by the questions, *"What will (s)he*

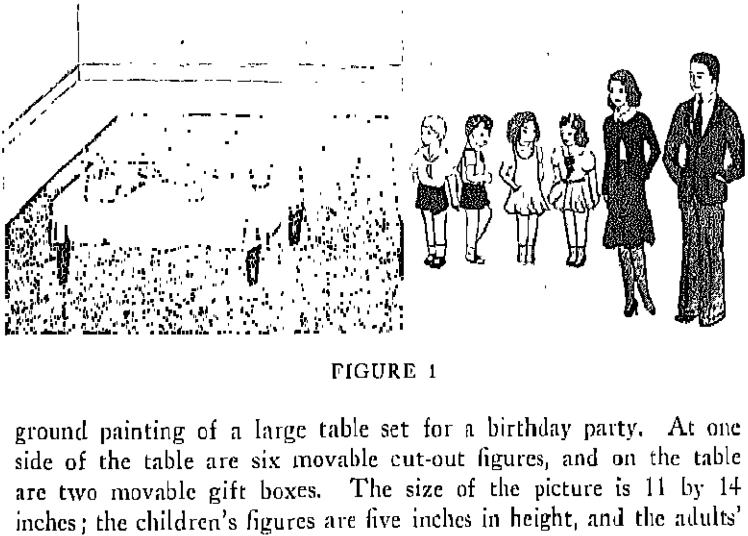

FIGURE 1

find in this box?" (pointing to the square box), and *"What will (s)he find in this box"* (pointing to the long box).

Figure 2 illustrates a picture (living room opening out-of-doors),

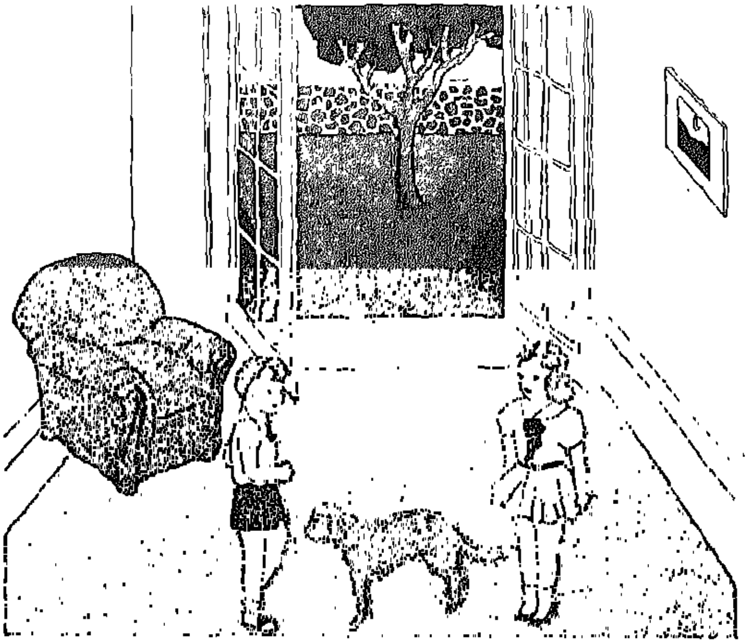

FIGURE 2

with three movable figures (boy, girl and dog); Figure 3 a child's bedroom opening into another bedroom, with moveable figures of a bed, children, and a ball, and of two adults placed beside the picture; and Figure 4 an outdoor playground with movable figures of children and a ball. Figures 2, 3 and 4 are introduced with the question, *"What is this picture about?"* and after this has been answered, *"Can you make another picture?"* Figure 5, the last in the series of colored background pictures with movable figures, is presented with the questions, *"Whose play-room shall we call this?"* and *"What does (s)he want from the shelves?"*

Figures 6 through 10 illustrate black and white silhouettes, 8¼ by 11 inches, and Figures 11 through 15 are photographs of an

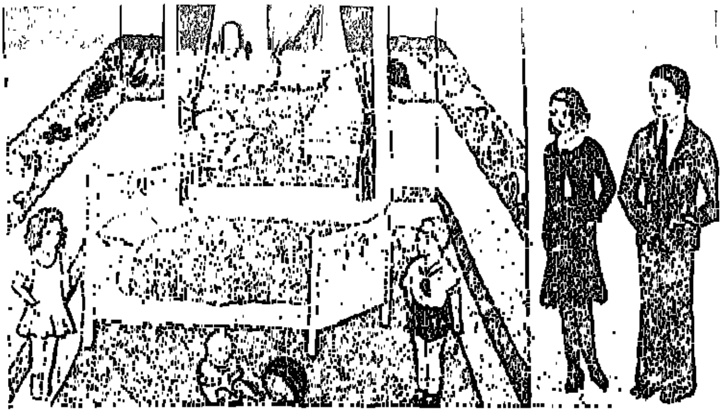

FIGURE 3

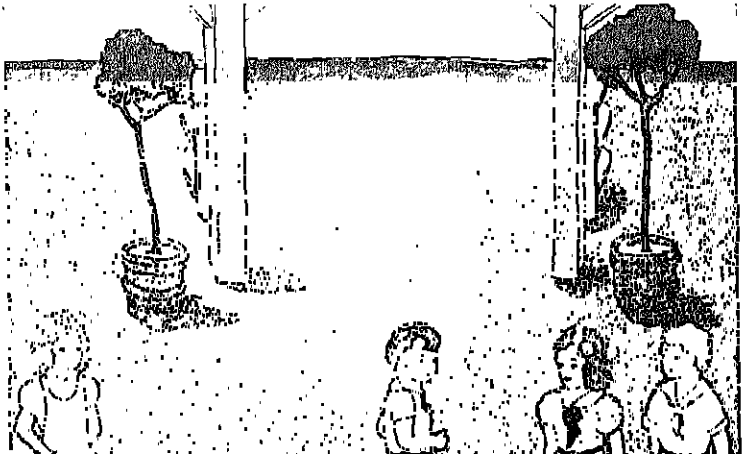

FIGURE 4

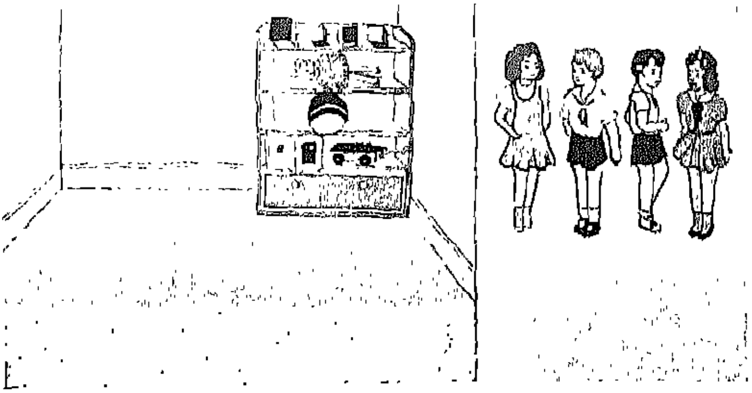

FIGURE 5

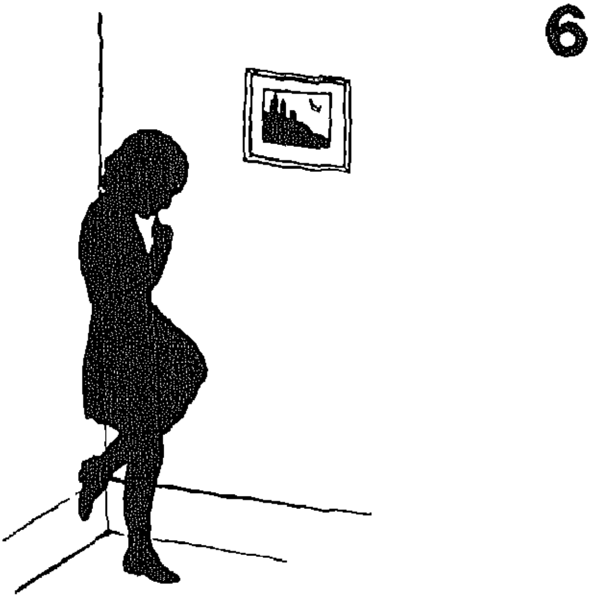

FIGURE 6

7

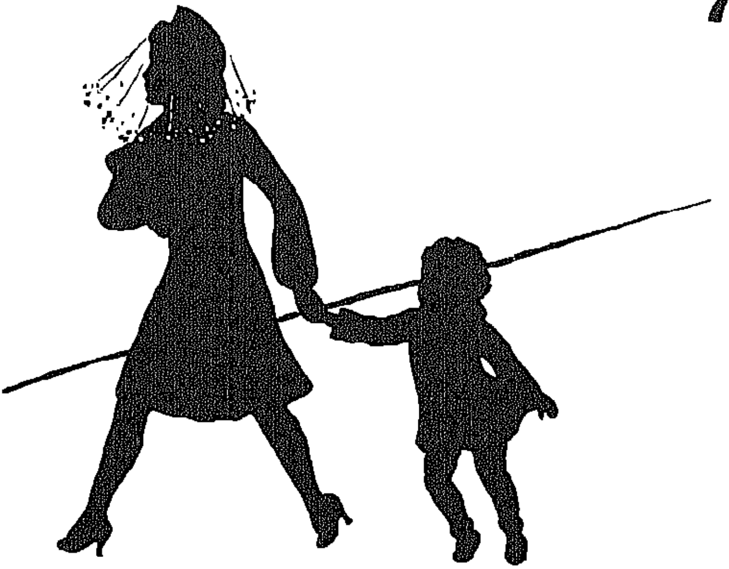

FIGURE 7

8

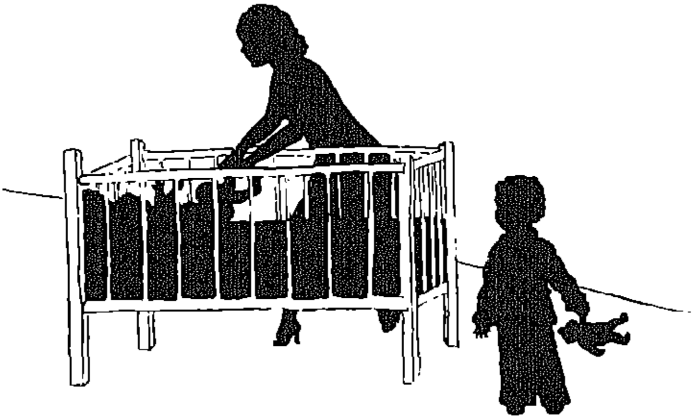

FIGURE 8

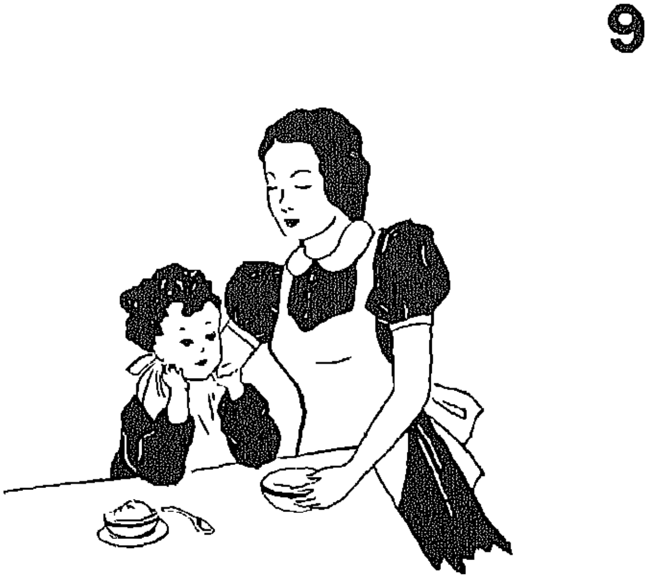

FIGURE 9

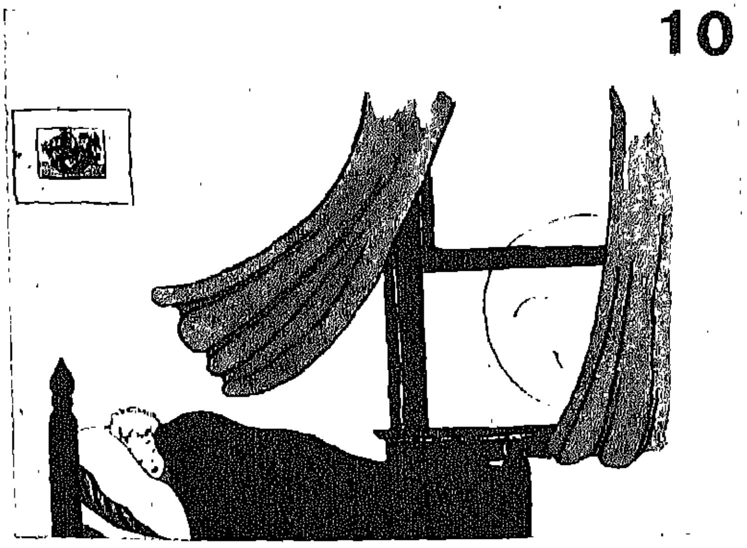

FIGURE 10

TABLE 1

CHRONOLOGICAL AGE OF THE SAMPLE

	Boys		Girls		Total group		
	N	Mean *CA*	*N*	Mean *CA*	*N*	Mean *CA*	Range
Two-year-olds	10	2 yr. 7 mo.	10	2 yr. 7 mo.	20[1]	2 yr. 7 mo.	2 yr. 3 mo. to 2 yr. 11 mo.
Three-year-olds	14	3 yr. 4 mo.	14	3 yr. 5 mo.	28[2]	3 yr. 4 mo.	3 yr. 0 mo. to 3 yr. 11 mo.
Four-year-olds	15	4 yr. 6 mo.	14	4 yr. 4 mo.	29[3]	4 yr. 5 mo.	4 yr. 0 mo. to 4 yr. 11 mo.

[1] Twelve from *ICW*, 8 from Whittier and Edison.
[2] Nineteen from *ICW*, 9 from Whittier and Edison.
[3] Eleven from *ICW*, 18 from Whittier and Edison.

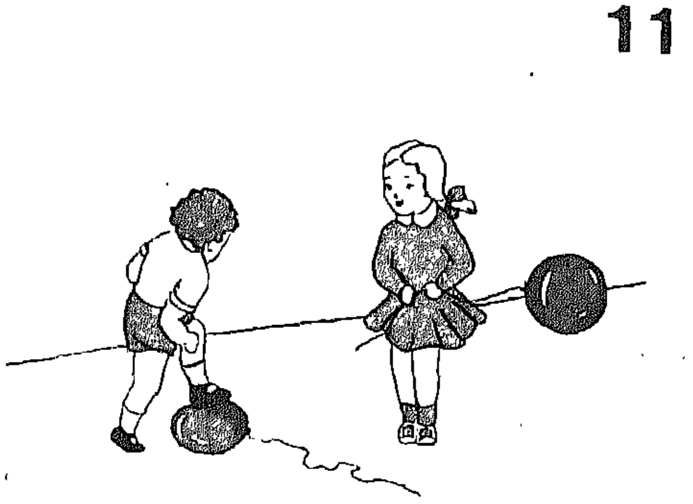

FIGURE 11

TABLE 2
MEAN INTELLIGENCE QUOTIENTS OF THE SAMPLE

	Boys		Girls		Total group		
	N	Mean *IQ*	*N*	Mean *IQ*	*N*	Mean *IQ*	Range
Two-year-olds	10	118	10	107	20	113	75 to 173
Three-year-olds	14	120.5	14	119.5	28	120	76 to 165
Four-year-olds	15	112	13	115	28	114	72 to 146

additional series of colored pictures of the same size as the silhouette. Each of these is introduced with one or more of the following questions, *"What is this a story about? . . . What does it make you think of? . . . Let's tell a story about the picture."*

C. SUBJECTS

The age range and distribution were as follows: 77 children, 39 boys and 38 girls, ranging in age from 2 years and 3 months to 4 years and 11 months, acted as subjects in the experiment. All of the

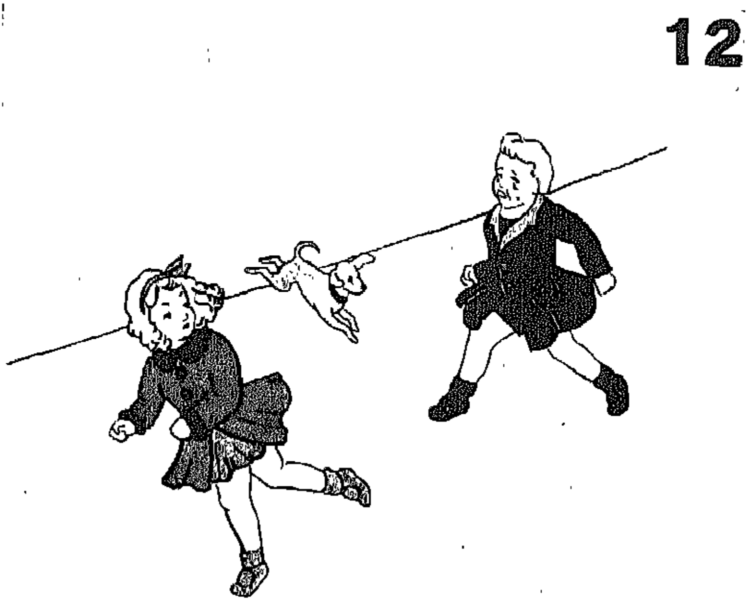

FIGURE 12

children belong to nursery school groups in Berkeley, California; 42 were from the Institute of Child Welfare nursery school; and 19 and 16 respectively from the Whittier and Edison nursery schools. The latter two were units of the emergency nursery school program operated by the Works Progress Administration.

Tables 1 and 2 give the characteristics of the sample with regard to *CA* and *IQ*.

The investigation deals with children of varied cultural backgrounds; the Institute of Child Welfare sample represents a superior environment in terms of the criteria most commonly used. The majority of families belonged to the "professional" or "business executive" brackets in the socio-economic scale. Among the children at the *WPA* nursery schools were five negroes and five Japanese. A careful study of their records showed no evidence of language difficulty or other handicap in responding to the pictures, and they were therefore retained in the sample.

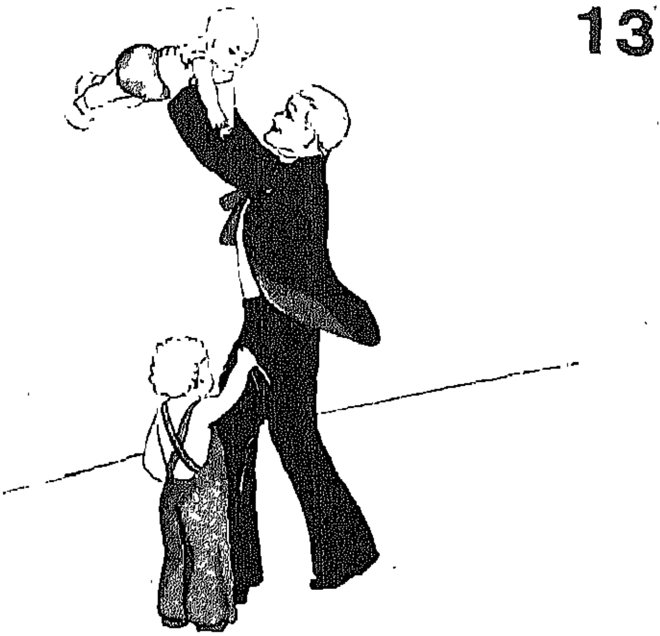

FIGURE 13

D. METHOD

The data for this experiment, the responses of 77 nursery school children to a series of 15 pictures, were collected during an eight weeks' period between March 27th and May 22nd, 1940; 1,155 responses were recorded; 17 of these were negative (refusals or failures to respond).

The colored background pictures with movable parts were shown first during an experimental period of approximately 15 minutes. The session was limited in time in order to avoid fatigue or loss of interest. The time element was not stressed, however, and if a child became absorbed in the material and was unwilling to leave at the end of the 15-minute period, additional time was allowed.

The remaining 10 pictures were shown in a second experimental period of 15 to 20 minutes. In one instance three experimental periods were required in order to complete the series.

Responses were recorded verbatim by the experimenter on record

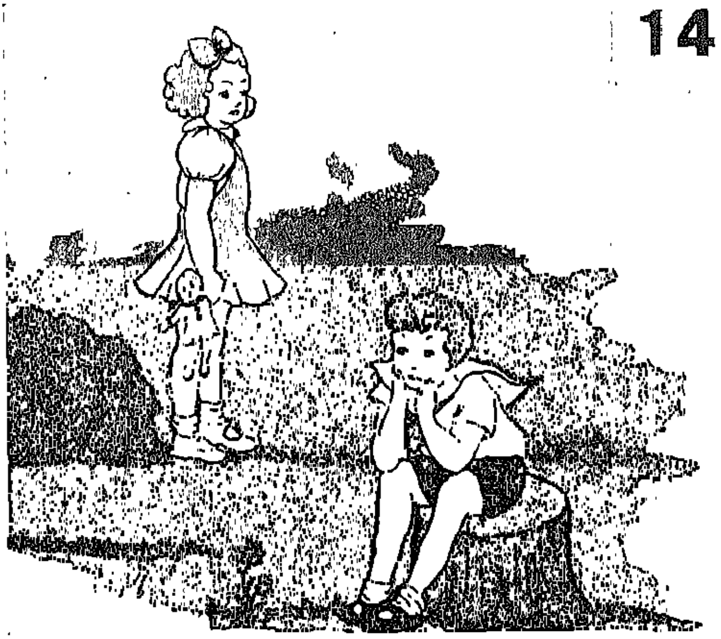

FIGURE 14

sheets prepared for this purpose. Diagrams on the first page of the record sheet reproduced the major features of the background pictures with movable parts, making possible a rapid recording of the children's reconstructions of these pictures.

The experimental periods were scheduled between 9 and 11:30 in the morning, except for a group of 14 children enrolled in an afternoon session of the Institute of Child Welfare Nursery School, University of California. An effort was made to keep the experimental situation as natural and as enjoyable as possible for the child of nursery school age. An attractive, quiet room was available for experimenting in each of the three schools whose children participated in the experiment. Unusually complete nursery school records were available for comparison with the projection records, and nursery school teachers were unfailingly coöperative.

The most essential feature of the methodology of this projection experiment was the use of experimental material which proved in-

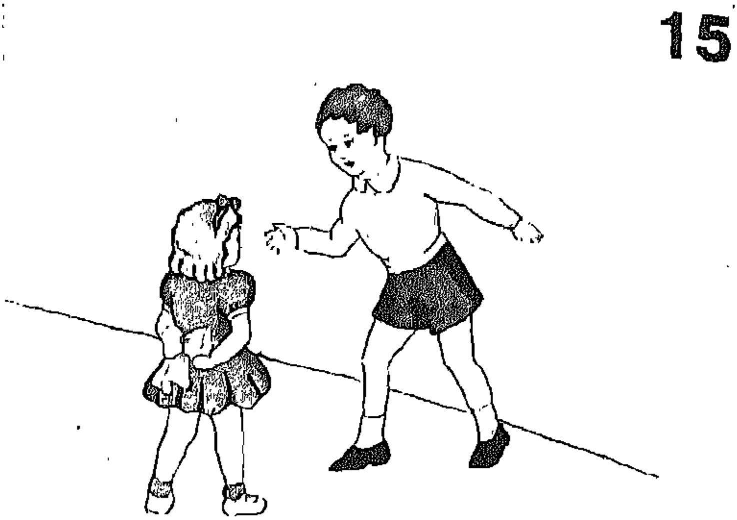

FIGURE 15

trinsically interesting to children of two, three, and four years of age, and which could be employed effectively in stimulating "personal" responses.

The following comments from individual records illustrate the favorable reception of these pictures by children at various age levels.

Judith, 2 years, 11 months. Picture 11:

> "Pretty picture! There's a girl and there's a ball. That's a girl and that's a ball."

Anne, 3 years. Picture 14:

> "Oh dis is so purty! Will you buy me some for my home, so my mudder can show them to me? White and brown and red and black! Dis is so purty!"

Donald, 3 years, 3 months. Picture 15:

> "They're talking, aren't they?" (End of experiment.) "I want to look at them all in a row. I like dat picture the best. (Balloons.) I like dat too the best. (Day dreaming.) Nice pictures! Maybe you'll show them to me again."

Peter, 3 years, 5 months. Picture 14:

"What's that? What is it? A big girl and a boy, isn't it? Have you got any more? Give me some more. Give me a lot of it!"

Carol, 3 years, 8 months. Picture 11

"They both have balloons and the boy is stepping on his. The girl's blowed away. I think I like this the best. I like this the best of all!"

Teddy, 4 years, 7 months. Picture 14:

"That boy's sitting on a little tree. That girl's got a doll—a green tree. That's green! I wish I could keep all these pictures!"

Bill, 4 years, 7 months. Picture 15:

"Oh, that's a pretty picture! Boy telling girl, 'You'd better go across the street.' She got run over too. Anything more you got for me to do?"

E. Illustrative Records

Before proceeding with the normative analysis, the reader is invited to examine illustrative records at each age level, for each of the pictures requiring a single non-manipulative response (pictures 6-15).

1. *Picture 6: "The Sulky Child"*

TWO-YEAR-OLDS

Marilyn: "A teacher."
Gerry: "Little dirl."
Margaret: "Man."
Ruth: "Tongue."
John: "Girl—candy."

THREE-YEAR-OLDS

David: "A girl walking along and the country." (*Pointing to picture.*)

Donald: "Girl—dancing."

Janet: "That's a girl going down the steps."

Roberta: "A little girl, going to school. She's playing around this thing."

Carolyn: "She's hopping." (*Points to picture.*) "A little moon."

FOUR-YEAR-OLDS

Sandra: "She thinks she's brushing her teeth."

Stephanie: "She standing there on one foot, thinking, isn't she?"

Roxanne: "A little girl looking at the picture. She thinks it's pretty."

Clifford: "That girl's all tired out and going to bed. That's Jocelyn. She goes to the big school."

Anne: "She's thinking.

2. *Picture 7: "Mother Coercing Child"*

TWO-YEAR-OLDS

Denis: "Boy—lady."

Delight: "A baby, a mama too—money."

Ruth: "Baby and mommie."

Nancy: "That's Ann." (*With great conviction.*)

Judy: "A lady and another."

THREE-YEAR-OLDS

Carol: "They're walking. The mother is walking with her baby."

Anne: "Mudder and a little boy and a stick going to the fair."

Gail: "I don't know. I think it's about Susie and her mother. I think they're going for a walk in the rain. See the drops!" (*Pointing to veil.*)

Sandy: "That lady's taking the little boy, and he's running. He has a mother and daddy. Tomorrow I'm going to get a new gun, but I'm not going to shoot anybody."

FOUR-YEAR-OLDS

Margaret: "Mother and a little girl. I guess they're going home. I guess they're going home because she is naughty."

Clifford: "That boy doesn't want to walk. That's a lady. She's going to spank the boy. He's a bad boy. He asks for something and he doesn't get nothing."

Helen: "Mother holding the baby's hand. She's (the baby) afraid of poking that in her eye."

Billie: "Looks like a woman and a little girl. Taking her to the hospital. I was born in the hospital in Oakland."

3. *Picture 8: "Mother Caring for Baby in Crib"*

TWO-YEAR-OLDS

Richard: "Lady and bed. Look it! Baby!" (*Pointing to little boy.*)

Nancy: "That bed my bed."

Marilyn: "A little bed. That dollie!" (*Pointing to toy.*)

Leon: "Little bed, mamma, baby."

Michael: "And here's anudder boy, and here's anudder mommie wif her hat off."

THREE-YEAR-OLDS

Gail: "What is it about? She's getting out of her bed. She's getting up. That's about getting up."

Bill: "There's the baby and here's the little boy, and the lady's fixing the baby and the baby's crying."

FOUR-YEAR-OLDS

Darleen: "What is Mommie doing in the bed? A baby! The little boy wants to get in the crib."

Clifford: "This little boy is out of his bed, and the mother tries to find the boy. This baby is in the boy's bed. The Mommie's a bad girl. That boy fights with this baby."

Betty: "That's a lady putting her baby in bed. She's crying. Look, she's kicking. She doesn't want to go to bed, does she?"

4. *Picture 9: "Child Refusing to Eat"*

TWO-YEAR-OLDS

Gerry: "Two dirls—dinner."

Margaret: "Milk."

Marilyn: "A girl—a mommie."

Kiashi: "Girl wiv a boy."

Shinobu: "Spoon down dere."

THREE-YEAR-OLDS

Karen: "A boy—he's tiny: not big. Mommie giving breakfast."

Roberta: "Oh, she doesn't have any feet, and neither does she! Now, you see, this little girl is eating her breakfast, and now the Mommie is giving the little girl her soup. My Mommie used to curl her hair. Now she rolls it up."

Gordon: "Eating supper. He eats potatoes."

Carlo: "What can they eat there? She's got her eyes closed. What for? What can the boy do?"

Michael: "Baby's having dinner. That's a dishbowl. Mommie eating her dinner. Dat all."

FOUR-YEAR-OLDS

Mike: "A little boy doesn't want to eat, and his mama's going to get him to."

Midorie: "Eating lunch, this little girl and she don't want to. She got curly hair."

Bobbie: "I see a little girl. Mama makes the little girl eat her dinner, but she doesn't."

Clifford: "That boy's going to eat up his soup. He's a bad boy. He won't eat his soup. He just sits there. He could eat the steak. That's my mommie. That's my baby. She's going to grow big like Jocelyn. She won't get a spanking, because her mommie's going away. She has to mind."

5. *Picture 10: "Child Sleeping"*

TWO-YEAR-OLDS

Margaret: "It's Mama."

John: "It's a girl—a bed."

Marilyn: "This girl."

Denis: "Pillow."

Kiashi: "Girl—sleepy."

THREE-YEAR-OLDS

Carol: "It's night and the moon came down, and the girl is sleeping."

David: "Her sleeping. Why sleeping now?"

Peter: "It's about a little girl, going to sleep."

Marilyn: "Dat's a little girl, sleeping in a bed under the covers."

Karen: "A lady in bed sleeping. The moon out there. The sunshine going down. Something up there." (*Pointing to picture.*)

FOUR-YEAR-OLDS

Mike: "A little boy asleep, but what's the sun laughing about? I know. The sun is out and the wind's blowing too, and that's why he's laughing."

Clifford: "A big moon looking in the window. He's in bed. The moon laughs funny. A bad moon."

Midorie: "That little girl is asleeping. What's this? Bad wolfie here. This bad wolfie can't see, because she close it. She got no curlie."

Billie: "Oh, I think this is about a boy sleeping, and the moon is coming in and the window is open, and the curtains are waving. The boy is sound asleep. I like sleeping. I sleep well!"

6. *Picture 11: "Balloon Play"*

TWO-YEAR-OLDS

Jimmie: "Dat a red one. Apple. Dat's green."

Gerry: "A little boy and a dirl. Now another!"

Judith: "Pretty picture! There's a girl and there's a ball. That's a girl" (*pointing to boy*) "and that's a ball."

Margaret: "Dirls."

Leon: "Balloon—balloon!"

THREE-YEAR-OLDS

Peter: "That's about a little girl and a little boy, playing outside."

Karen: "The boy jumping on his balloon and pop it. They have balloons. A tree lying down there." (*Pointing to floor line.*)

Anne: "Oh, look at those, that and that and they're playing. Boy and girl."

FOUR-YEAR-OLDS

Stephanie: "He's stepping on his balloon and popping it. She's holding hers so it doesn't pop. I guess he doesn't want a balloon. I kind of like things colored."

Clifford: "That boy trying to pop his balloon. He wants to make a kite out of it, but he can't when he's popping it. That would be a kite." (*Pointing to girl's balloon.*) "What the girl wants is a kite. She wants to make a kite out of it. He may step on it, but it won't hurt me."

Craig: "Dat the balloon! Dat boy popping his balloon. Dat girl don't want to pop her balloon!"

Marilyn: "Balloons! Look what he's doing! The girl's not naughty. He's naughty. He's got baby's shoes! That's purple." (*Pointing to girl's dress.*) "Purple socks. What's that?" (*Color, pointing to red balloon.*)

Bobbie: "Dey little boy, trying to pop his balloon, and the little girl isn't popping hers. She's just holding hers. He's popping it. He doesn't want it."

7. *Picture 12: "Two Children with Dog"*

TWO-YEAR-OLDS

Judy: "That's a picture of that boy and that girl and that doggie and that line."

Denis: "Girl and a boy and a doggie."

Ruth: "Doggie—high-doggie!"

Shinobu: "Boy, puppy, girl."

John: "It's a girl. What on face? Funny!"

THREE-YEAR-OLDS

Carolyn: "A dog running and a little boy and little girl running. A little dog and boy and girl running."

Bill: "Oh, the dog's jumping over the fence, and the girl's running away from the dog."

Anne B: "Oh, dey's all running away! Aren't dey funny? Dey's funny! Dey's all funny!"

FOUR-YEAR-OLDS

Mike: "They're running along because they don't want the dog. He's crying a little. He's found his dog, so I should think he would be happy."

Anne H: "They're both running places. The little boy is sad, and the little girl is happy."

Sandra: "Crying, see" (*pointing to tears*) "because that little girl did something. And the little girl is happy."

Jonathan: "A girl and a bad man trying to hurt the girl and a doggie, see! He's going to do something to the bad man."

Dianne: "Little doggie on the line might bite the boy."

8. *Picture 13: "Father Tossing Baby"*

TWO-YEAR-OLDS

James: "One a boy." (*Pointing to older child.*) "Two a boy." (*Pointing to baby.*)

Judith: "That little girl" (*pointing to baby*) "and a little girl. Little blue pants on her and little red pants on her. Her hand! Her hand!" (*Pointing.*)

Marilyn: "Dat man. Baby way up in air!"

Margaret: "Girl way up high. Yeah-um-hum."

Delight: "He's a baby. My baby! Mine like dat. My baby way down dere—girl." (*Pointing to older child.*)

THREE-YEAR-OLDS

Gail: "Tossing! The father has the baby like this" (*imitates*) "and he's holding on to the skirt. We won't look at that any more."

Peter: "Taking a sunbath, this boy! What's that Daddy doing? Throwing that baby down on the ground!"

Anne: "Two little girls and a Daddy. Look what he's doing to the baby. My daddy does that to me sometimes."

Donald: "Oh, he's carrying him, the little baby! There's sunshine on dat one, isn't there?"

Karen: "A man putting tiny boy up in air. A girlie by him. Watching boy way up in air. A tree is by him." (*Pointing to floor line.*)

FOUR-YEAR-OLDS

Mike: "A papa lifting up the baby, and the little girl doesn't want him to. She's afraid he'll drop the baby."

Clifford: "This baby" (*pointing to older child*) "wants to get up here too."

Midorie: "Father's taking the baby. This little girl, she's saying, 'Father!' She wants to go for a walk."

Stephanie: "The father is holding the baby up and that one wants to catch the baby."

Bert: "Oh, man! A little girl and a Daddy pick up the baby. Hey, what's the girl doing to her Daddy? Trying to take the Daddy's money!"

9. *Picture 14: "Daydreaming"*

TWO-YEAR-OLDS

Richard: "Look, a boy on a dump. Look, a big girl and a dollie! Red!"

Margaret: "Girl, girl, dollie—yeah, more!"

Marilyn: "Little boy—a yndy."

John: "It's a boy! That little baby got shoes." (*Pointing to doll.*)

Kiashi: "Him, dat." (*Pointing to doll.*) "Him on dat."

THREE-YEAR-OLDS

David: "There's a little boy sitting on a rock, and a little girl, playing with a doll."

Donald: "Oh, having a good time out in the sunshine, aren't they?"

Keir: "Jumping!"

Gale: "He's sitting on something, isn't he? He's sitting on some wood. She's walking with a little girl. This is a doll."

Karen: "A boy sitting on a stump. A tree there. A house around there. A little dollie. A boy and a gallie playing with a little dollie."

FOUR-YEAR-OLDS

John: "A boy thinking and here's a little girl with a doll in her hand and here's a green hill."

Phil: "Sitting on a tree. He's putting his hand on his face, because he's afraid he's going to fall down."

Clifford: "That's a little boy sitting down on a stone. She's holding her dollie. She got lost. They got lost. The boy will get a gun and shoot both hands off. A boy in school had a gun and shot these off." (*Pointing to fingers.*) "The cops could come. Another man shot a gun and shot his hands."

Roxanne: "A little boy thinking. A tree chopped down. A little girl with a dollie. She has a pretty ribbon. A flower! Mama didn't plant them. God did it. They're very pretty."

Tommie: "What's that sad face? A tree stump! I don't like that dumb, dumb picture—a sad picture."

10. *Picture 15: "Big Boy and Little Girl"*

TWO-YEAR-OLDS

Nancy: "Girl and boy."

Judy: "That's a lady; that's a lady too, and that's a black line."

Denis: "A dress." (*Pointing to girl.*) "Pants." (*Pointing to boy.*)

John: "Boy and that's a girl."

Gerry: "A boy and dirl."

THREE-YEAR-OLDS

Keir: "He's jumping. Oh, he's naughty. Someone take his hands away!"

Gail: "They're playing. They're not doing anything."

Allen: "It's a boy and a girl doing that."

Bill: "The girl's standing and the boy's walking."

Karen: "A little gallie and boy playing togedder. The boy is bigger'n the gallie. They're playing with nothing. A tree lying there." (*Pointing to base line.*)

FOUR-YEAR-OLDS

Billie: "Oh, that's a pretty picture! Boy telling girl, 'You'd better go across the street.' She got run over too."

Lynette: "He says, 'What hand do you want?' She's playing."

Clifford: "Oh, that's a little boy and a little girl. The boy's going to pick her up and take her to her mother."

John: "That's a boy and he's trying to fight the little girl."

Marilyn: "They're speaking to each other. I think they're going to a birthday party."

III. ANALYSIS OF DATA

A. Categories of Perceptual Response

In the foregoing examples three major types of response can be observed: (*a*) A simple naming or other identification of objects. This may be regarded as response in terms of static form or enumeration (*"A boy, a lady"*). (*b*) The description of the picture situation in terms of overt activity (*"This little girl is eating her breakfast"*). (*c*) Inference as to psychological states or inner activity (*"A little boy doesn't want to eat and his mama's going to get him to"*).

When we study these categories of response, with reference to age level, it is apparent, from the examples given, that the two-year-olds respond predominantly in terms of static form. Description in terms of overt activity is more common after three years, and the suggestion of inner activity, rarely shown at two years, is fairly common at four years. To some extent these differences must be imputed to differences in language ability, but the protocols give fairly direct evidence, which will be considered in a later section, of age differences in the content of what is observed and in the orientation of interest. Table 3 gives the percentages for each picture, by category and by age group. The various pictures yield similar results on the whole: descriptions in terms of static form range from 58 per cent to 88 per cent in the two-year group; and from 7 to 46 per cent in the four-year group; interpretations in terms of inner activity range from 0 to 21 per cent in the three-year group; and from 11 to 41 per cent in the four-year group. Table 4 presents the group means.

Two age-trends are therefore clearly discernible in the analysis of the records of response, an increase with age of the tendency to interpret picture content in terms of activity rather than in terms of enumeration or static form, and an increase also in the tendency to interpret pictures in terms of inner rather than of outer activity.

It is probable that the age differences between the three- and four-year groups are somewhat reduced by the fact that the four-year-olds contain a larger proportion of children from the emergency nursery schools and have a somewhat lower mean *IQ* than the three-year-olds.

Within each group a relation also exists between *IQ* and scores derived from the percentage of pictures in which an individual em-

TABLE 3

CHILDREN'S RESPONSES TO PICTURES: PERCENTAGES IN EACH OF THREE MAJOR CATEGORIES

Pictures		Static form			Outer activity			Inner activity	
	2 years ($N = 20$) %	3 years ($N = 28$) %	4 years ($N = 29$) %	2 years ($N = 20$) %	3 years ($N = 28$) %	4 years ($N = 29$) %	2 years ($N = 20$) %	3 years ($N = 28$) %	4 years ($N = 29$) %
No. 6	88	70	46	12	30	36	0	0	18
No. 7	84	41	21	16	56	41	0	5	38
No. 8	70	32	21	30	57	62	0	11	17
No. 9	58	30	7	42	55	61	0	15	32
No. 10	63	29	25	32	61	64	5	10	11
No. 11	80	21	14	20	57	52	0	21	34
No. 12	65	36	10	35	46	48	0	18	41
No. 13	80	36	24	20	43	52	0	21	24
No. 14	65	39	21	35	50	62	0	11	17
No. 15	80	44	43	20	44	36	0	11	21

TABLE 4
MEAN PERCENTAGES OF RESPONSES, PICTURES 6-15

	2 years	3 years	4 years
Static form	73	38	23
Outer activity	26	50	51
Inner activity	1	12	25

ploys a description or interpretation of activity. Product-moment correlations are .51 for the two-year-old group; .68 for the three-year group; .50 for the four-year group. It may be noted that the three age groups give higher product-moment correlations when the percentage of interpretation in terms of activity is compared with *IQ* than when it is compared with *CA*. In the latter case, the correlations for the three age groups are, respectively, —.13, .53, and .40.

B. MISINTERPRETATION OF DETAILS

Another striking age-trend in children's reactions to pictures is from perception of concrete, often unrelated details of a situation, to a recognition of these concrete details as parts of some larger whole. This is illustrated in the following examples selected from picture descriptions at different age levels. Misinterpretations due to over-literal, concrete observations are italicized.

TWO-YEAR-OLD RESPONSES

Michael. Picture 3:

"Dat a *tunnel.*" (*Pointing to open space in bureau of smaller room.*)

Jimmie. Picture 3:

"These are *button-holes.*" (*Pointing to open spaces under Daddy doll's bent arm.*)

Michael. Picture 10:

"Dat mommie's asleep in bed. Look at the *dress teared.*" (*Pointing to curtains.*)

Richard. Picture 12:

"Doggie, doggie, man, girl. *Doggie jumping the rope.*" (*Pointing to floor line.*) "I jump."

THREE-YEAR-OLD RESPONSES

Bill. Picture 3:

Playing with ball, "It lands on the baby's nose. Peek-a-boo! She threw it in the *water.*" (*Pointing to blue floor.*)

> (*Observing small objects in a blue decorative wall border.*) "Hi, a boat sailing here! Do chickens have to be in the *water?*"
>
> *Karen. Picture 3:*
>
> "There baby's bed and there a chicken—there the *water.*" (*Pointing to blue floor.*)
>
> *Carlo. Picture 3:*
>
> . . . "What can the boy do? He has only *one foot.*" (*Pointing to profile figure.*)
>
> *Gail. Picture 7:*
>
> "I don't know. I think it's about Susie and her mother. I think they're going for a walk in the *rain. See the drops.*" (*Pointing to veil.*)
>
> FOUR-YEAR-OLD RESPONSES
>
> *Bobby. Picture 3:*
>
> "A mother and a daddy, and a little baby and a little girl and a little boy and *a pond.*" (*Pointing to blue floor.*) "A starfish and a rooster."
>
> *Jonathon. Picture 11:*
>
> "A girl and she's by a ball and a boy. He's stepping on a *big apple with a great big stem on.*"
>
> *Midorie. Picture 12:*
>
> "That doggie! Another girl! That girl." (*Pointing to boy.*) "Got all *dirty.*" (*Pointing to tears.*)
>
> *Marilyn. Picture 13:*
>
> "The baby and the Daddy and the boy *going down the hill.*" (*Pointing to floor line.*) "Both ways they are. Look at the baby! Isn't she big?"

From examples of the preceding type it is apparent that children in all three age groups frequently misinterpret picture content through interest in detail, which has not been accurately related to a whole of which it is a part.

In contrast to the general tendency toward concrete rather than symbolic interpretation of pictures throughout the preschool period of development, some four-year-olds give evidence of the capacity to interpret an objective aspect of a picture, such as posture in Picture 6, in terms symbolic of thought or feeling. Whereas there are only two references to general posture among the 20 two-year-old responses to Picture 6, *"Jumping"* and *"Leaning down,"* and many references to posture in the three-year-old group, as *"Looking," "Putting finger in mouth," "Walking along," "Dancing," "Going down the steps," "Going to school," "Playing," "Hopping,"* there are five four-year-old children, three belonging to the highest quartile for intelligence level as measured by the Stanford-Binet Scale, who interpret the posture of the child in this silhouette as a symbol of thinking or feeling. Furthermore, no two-year-old has understood the symbolism expressed in Picture 7, coercion. One three-year-old interprets the objective details of posture in this picture symbolically as follows:

Carlo:

> "She's a hat on. What can the boy do? Where's the mama going to take the boy? Have they been naughty? I think they have been naughty. She has a hat on. What has he got on?"

Four of the four-year-olds interpret the figures symbolically as expressing coercion of the child by the mother.

Margaret:

> "Mother and a little girl. I guess they're going home. I guess they're going home because she is naughty."

Anne:

> "It's about a mother taking her down to have her hair cut."

Billie:

> "Looks like a woman and a little girl. Taking her to the hospital. I was born in the hospital in Oakland."

Clifford:

> "That boy doesn't want to walk. That's a lady. She's going to spank the boy. He's a bad boy. He asks for something and he doesn't get nothing."

In Picture 9, *Child Refusing to Eat,* no two-year-old interprets the child's bodily attitude as symbolic of refusal to eat. Two three-year-olds explain the bodily posture as follows:

Gale:

> "Shall he eat this or shall he eat this?"

David:

> "A little boy. He's going to eat. That's too hot for him."

Eight four-year-old children interpret the child's bodily attitude symbolically as resistance to eating.

Mike:

"A little boy doesn't want to eat, and his mama's going to get him to."

Anne:

"He's eating his dinner, isn't he? He's just getting ready to."

John:

"A girl is trying to sit him down because he won't eat nicely. The little boy won't eat."

Helen:

"The baby is eating her food. Another one! She's thinking of eating her food."

Stephanie:

"She's supposed to be making him eat."

Bobbie:

"I see a little girl. Mama makes the little girl eat her dinner, but she doesn't."

Clifford:

"That boy's going to eat up his soup. He's a bad boy. He won't eat his soup. He just sits there. He could eat the steak."

Midorie:

"Eating lunch, this little girl, and she don't want to. She got curly hair."

The responses to Picture 12, *Two Children with Dog,*[9] illustrate age differences in reactions to a picture of distress. Two of the 20 two-year-old children comment on the boy's tears:

James:

"What dat boy doing? Crying. What's he got on dose legs? What dat dog doing? What dat boy doing to dat dog?"

John:

"It's a girl. What on face? Funny!"

Four of the 28 three-year-olds comment on the boy's tears.

David:

"A little boy and a little girl playing with a dog and the little boy is crying."

Michael:

"Crying." (*Pointing to boy.*) "Dere a doggie and a girl, dat's all."

Anita:

"Little boy cry. What for? What for he cry?"

Peter:

"A cow and a girl and a boy. I think they're running, don't you? I think he's crying because he wants to go home."

Eleven of the 29 four-year-olds comment on the boy's tears.

Mike:

"They're running along because they don't want the dog. He's crying a little. He's found his dog, so I should think he would be happy."

Anne:

"They're both running places. The little boy is sad, and the little girl is happy!"

Sandra:

"Crying, see" (*pointing to tears*) "because that little girl did something. And the little girl is happy."

Stephanie:

"Oh, she's running. She socked him, and he's going to sock her back, and the dog comes along."

Roxanne:

"About a little boy crying, and a little dog running. That's why he cries. A girl has a green coat and a ribbon to match her coat, and socks to match her ribbon and dress. I got a turquoise coat for Easter. You mustn't wear it on rainy days. It will spoil it."

Lynette:

"The morning! Him crying. Dat dog walking. Putting on dem coat." (*Pointing to girl.*)

Bobbie:

"They're running somewhere and the little dog's running after. The little boy's crying because the little girl's running faster—running down the street."

Marilyn:

"Dog and a boy and that girl. He's crying because that dog. I could run fast. Where they going? What's that dog for?

> Maybe he doesn't know where the ball is. Are they going up the hill? When it's upside down, they're going up the hill."
>
> *Philip:*
>
> "Now they're running. The dog is running after these. This is crying."
>
> *Clifford:*
>
> "The girl runs as fast as she could. The boy cries. The boy cries because of the dog. He's afraid. I'm not afraid of dogs. I have a little dog. He's white and black."
>
> *Teddy:*
>
> "That girl hit that little boy and he cried, and that dog came running fast. That boy can't run fast and the girl can. That boy has a fat head."

The conclusion to be drawn from the above illustrations, which are fairly typical of the responses to all the pictures, seems an obvious one; namely, that young children's interpretations of pictures tend to be very objective and concrete. Capacity to interpret objective details not only as part of a larger whole, but as symbols of subjective experience under appropriate circumstances, represents a maturity of response which is reached infrequently by the two-year-old children acting as subjects of this experiment, somewhat more frequently by the three-year-old children, and considerably more frequently by the four -year-olds.

C. PART-WHOLE BALANCE IN PERCEPTION

A third and clearly-demonstrable age-tendency, observable in a study of the data of this experiment, is a sequence of development with reference to part-whole balance in perception. The stages of development in perception with reference to part-whole balance appear to be the following, starting with the least mature type of perceptual response given in the records under consideration: (*a*) A part or detail interpreted as a whole. In the most immature examples of perception, the detail reported as a whole is frequently an insignificant or irrelevant detail. (*b*) An unanalyzed but relatively complete whole. (*c*) An incomplete whole, analyzed incompletely. (*d*) A whole in which some detail is discriminated. (*e*) A whole rich in its details.

The responses of all three age-groups to three of the pictures calling for verbal interpretation are summarized below. The pictures selected for this purpose are Picture 6, *The Sulky Child*, selected because it is the more difficult of two pictures containing a single, central figure; Picture 8, *Mother Caring for Baby in Crib, with Older Child Looking On;* and Picture 13, *Father Tossing Baby, with Older Child Looking On;* selected because they are the only pictures in the series from 6 to 15, containing three central figures.

Good approximation in report to the stimulus presented may be used as the criterion of mature perception. For this reason, meaningful material has a great advantage over the standard inkblot materials in the study of part-whole balance in perception.

Responses to Picture 6 are summarized in Table 5. One of the difficulties of reporting projective data in quantitative terms becomes apparent in a summary of this type. Although the major categories of analysis stand out clearly, there are qualitative differences which cannot be taken account of in the above type of scoring. For example, under the major categories of analysis, no differentiation seems possible in quantitative terms of the type of whole or of the type of detail occurring in the individual record. Records of the two-year-old children list as the unanalyzed complete whole such diverse interpretations as "girl," "teacher," "man," "lady." A two-year-old interpretation of the picture as "tongue" represents a less comprehensive detail for carrying the meaning of the whole than the four-year-old interpretation, "*I think that's a boat,*" pointing to arrangement of floor lines.

TABLE 5

RESPONSES TO PICTURE 6

	2 years ($N = 15$) %	3 years ($N = 26$) %	4 years ($N = 28$) %
Detail interpreted as whole	7	0	3
Unanalyzed complete whole	73	35	11
Whole qualified by one detail	20	34	53
Whole qualified by two or more details	0	31	33

Responses to Pictures 8 and 13 are summarized in Table 6. While Table 6 reveals a consistent increase with age in the amount of qualifying detail reported, we must again point out the difficulty of stating qualitative differences in quantitative terms. The follow-

TABLE 6

RESPONSES TO PICTURES 8 AND 13

	Picture 8			Picture 13		
	2 years ($N = 20$) %	3 years ($N = 27$) %	4 years ($N = 28$) %	2 years ($N = 20$) %	3 years ($N = 28$) %	4 years ($N = 29$) %
Detail interpreted as whole	10	7	0	5	4	0
Whole, including one major character	20	11	4	20	4	3
Whole, including two major characters	65	29	32	55	25	24
Whole, including three major characters	5	53	64	20	67	73
One qualifying detail	40	39	50	40	18	28
Two or more qualifying details	10	32	32	20	68	69

ing response of a two-year-old child to Picture 13, *"Oh in the sky, in the sky,"* is qualitatively better than the response of a three-year-old child to the same picture, *"That is red trousers on,"* though both represent a very immature perception.

The three dominant age-trends already discussed, i.e., (*a*) the developmental sequence of three major patterns of interpretation, static form, outward activity, inner activity, (*b*) the trend of development from literal to symbolic interpretation, and (*c*) the five progressive developmental changes in part-whole balance in perception, will be useful background against which to interpret the individual reports of children in different age groups.

D. FICTION PLAY

1. *The Sex of Preferred Play Objects*

Several of the pictures (notably Nos. 1 and 5) were devised so that each child could make a choice of one or more figures represented: of children or adults, and of the same or opposite sex. In the response to Picture 1, as an example, many of the two-year-old children immediately became absorbed in the picture-situation, unable to maintain the psychic distance necessary for fiction play. The two-year-old child's tendency to project himself into the picture-situation, without formal distinction between picture set and external reality, gave many excellent illustrations of what Piaget (6, p. 377) has aptly called the naïve realism of little children.

Four of the 20 two-year-olds, instead of selecting one of the movable figures for the birthday honors, responded as follows:

Richard. Picture 1:

"My!" Dickie's!"

Marilyn. Picture 1:

"Yeah-my!"

Denis. Picture 1:

"Me."

John. Picture 1:

"Mine."

Eight of the two-year-old group made no choice, largely through inability to comprehend the fiction play involved. The attitude was one of willing acceptance of a situation not understood, as:

Margaret. Picture 1:

"Ya—ya."

Marion. Picture 1:

"All wight—allwight."

Margaret K. Picture 1:

"Birthday."

In general, acquiescence in the play-situation and interested manipulation of the movable parts of the picture, give the typical pattern of response for this group.

Of the eight two-year-olds entering into the fiction-play, five made a choice, suggesting an easier extension of interest for, and a more sympathetic self-identification with dolls of the same sex. In the three-year-old group of 26 children, 73 per cent chose dolls of their own sex. In the four-year-old group including 29 children, 76 per cent of the group selected a doll of their own sex.

In the case of Picture 5, the question, *"Whose play-room shall we call this?"* brought out the same general type of response from the two-year-old group as was given to Picture 1, *The Birthday Table.*

Again 4 of the two-year-olds reacted as follows:

Richard. Picture 5:

"My."

Jimmie. Picture 5:

"I want it."

Marilyn. Picture 5:

"I will have it."

John. Picture 5:

"Will be mine."

Seventy per cent of the two-year-old group responded with instant manipulation of the toys, but failed to make a choice among the paper-dolls of an owner of the play-room and the toys.

Of the 26 three-year-olds who made a choice of the owner of the play-room, 62 per cent selected a doll of the same sex. Five girls and two boys gave joint ownership to all four dolls. One boy selected a girl and boy as joint owners. Two boys projected themselves into the picture-situation, as follows:

Donald. Picture 5:

"My room! I have my room at home!"

Bill. Picture 5:

"Me. Somebody else can come to my birthday."

Of the 29 four-year-olds, 72 per cent chose a doll of the same sex for owner of the play-room. Others gave the following responses:

Donald. Picture 5:

"All of them."

Betty. Picture 5:

"These little girls and boys."

Darleen. Picture 5:

"I would like to have a boy and a girl."

Craig. Picture 5:

"Craig's! These are my blocks, and this and this is! All these!"

Picture 6, *Coercion*, contains the figure of a child which can be interpreted either as a boy or as a girl. In the total group, the figure was described as a boy by 77 per cent of the boys and by only 16 per cent of the girls. With this picture, same-sex choices were less clear for girls than for boys, since many of the girls referred to the child merely as *"a baby."* Combining the results for Pictures 7, 8, 9, 10, 13 (all containing possibly equivocal figures) a total of 270 interpretations in terms of sex were obtained; 67 per cent of these were in terms of own sex, 33 per cent in terms of opposite sex. No definite age trend could be established, in the proportion of same-sex and opposite-sex choices. The results obtained are in general consonant with the belief that in story construction a factor of some importance is the identification of the story teller with some person in the story.

2. *Animal vs. Human Interest*

Picture 2, the living-room opening out-of-doors, was presented to the children with a little boy and girl standing in the center of the living-room, and a dog between them. Through the question, *"What is this picture about?"* and through the suggestion for rearrangement of the movable parts, opportunity was given for comparing the strength of human interest with the strength of interest in animals, and for observing the characteristics of the human interest in each age group.

Two points of contrast stand out clearly in the reactions of children to this picture: the increase in human or social interest with age, and the increase with age in the amount of reconstruction of the picture, accompanied by fiction play. Table 7 summarizes the two points of greatest contrast.

TABLE 7

AGE CHANGES IN FICTION PLAY (PICTURE 2)

Limitation of description to comments on dog			Attempted reconstructions with fiction play		
Age	*N*	Per cent of group	Age	*N*	Per cent of group
2 years	20	40	2 years	20	20
3 years	28	25	3 years	28	64
4 years	29	3	4 years	29	93

The following are characteristic descriptions of Picture 2, selected from the reports of the two-year-old group:

Jimmie:

> "Dat a dog. He walks. He says, 'Bow-wow, wuf-wuf.' No eye dere." (*Puzzled by dog's profile.*)

Judy:

> "He's a cow." (*Lifting up the dog.*) "If cows are too close, they look like doggies."

Marilyn:

> "The doggie—wuf-wuf!"

Denis:

> "Doggie—I want that doggie. He's broken."

Typical reports of three-year-old children comment on the boy and girl, as well as on the dog.

Roberta:

> "Boy against the chair and the doggie's here and so is the sister."

Kier:

> "About boys and girls. That's a boy and that's a girl and that's a dog."

Karen:

> "There a tree and there a tree and there a door. There a boy and a girl and a dog."

Gale:

> "She goes to the dog and he's here." (*Pointing to dog.*) "It's a police dog, and he goes for anything. This is his chair." (*Pointing to blue chair.*)

Typical descriptions from the four-year-old group are the following:

Mike:

> "A little boy, looking at dog. Little girl is looking at the

> dog. All looking out the window, trying to go out to play."
>
> *Marilyn:*
>
> "Chair, blue chair, dog, boy, girl." (*Points to picture on wall.*) (*Looking at blue chair again.*) "I wish it was yellow! We have a yellow chair at home. I have a red coat."
>
> *Midorie:*
>
> "A house; you open doors. And this boy sits down with her. Doggie goes outside."
>
> *Darleen:*
>
> "That's a dog, a little girl and a little boy. Trees outside. Hey! Look at the boat." (*Pointing to moon in picture on wall.*) "Are there any more little girls coming in? Have you any little girl?"

A comparison of the descriptive comments, as well as the summary of reports, indicates an age change in interest in human relationships.

Among the four-year-olds, 93 per cent are concerned with fiction play centering about the boy and girl, or including the boy and girl as well as the dog.

The following examples suggest the type of response given by four-year-old children to Picture 2:

> *Bert:*
>
> "Oh man! A dog and boy and girl!" (*While rearranging picture.*) "A boy is hiding behind the Chesterfield from his sister. A dog is sitting on the Chesterfield."
>
> *Dianne:*
>
> "They would like the chair. There are two." (*Putting boy and girl together in blue chair.*) "I could take the doggie home with me."
>
> *Clifford:*
>
> (*Commenting while rearranging the picture.*) "Girl's outside and the dog is going to jump over the fence. He looks a bad boy. He's sitting down on the chair."
>
> *Craig:*
>
> (*Suiting the action to the word.*) "And right here is where the boy sits!"

The rearrangements of the pictures with movable parts were for the most part quite individualistic and in many instances gave revealing insights into the inner experience of the individual child. Examples of such individualistic treatment of the experimental materials will be given in the discussion of individual protocols.

In contrast to the individual treatment of the rearrangements of Picture 2, it was interesting to observe that seven four-year-old boys placed the boy doll in the comfortable arm-chair, and seven four-year-old girls placed the girl doll in the big chair.

There were eight instances in this group of children, as against the 14 cases listed above, where a boy, in his reorganization of the picture, gave the chair to a girl, or a girl to a boy doll. Three children allowed a boy and girl to share the blue chair. In 50 per cent of the rearrangements of this picture by four-year-old children, the dog was placed out-of-doors.

3. *Age Differences in Fiction Play*

Picture 3, the large bedroom opening into a smaller bedroom, when presented to the children, contained the following movable pieces: a large double bed in one corner of the room, a little girl standing near the head of the bed, a baby playing with a ball near the foot of the bed, a little boy nearby, and across the room the mother and father dolls. Rearrangements of this picture made it possible for each child to create family constellations or groups according to individual choice. Individual choice rather than group tendency seemed to prevail in the responses to this picture. Some of the individual choices will be discussed in a later section of this report.

Definitely preferred arrangements occur in 35 per cent of the reports of the two-year-old group, in 79 per cent of the reports of the three-year-old group, and in 93 per cent of the reports of the four-year-old group. The most frequently preferred rearrangements of the picture center about the following groupings, listed in order of frequency:

Mother and Daddy in the big bed.
Daddy and boy in one bed, and Mother and girl in the other.
Girl, boy, and baby in one bed.
Girl and boy in one bed.

Picture 4, the playground scene, when presented first, had in its foreground two little boys and one little girl playing ball together. Another little girl was placed outside the group at the extreme edge of the playground. The picture-reorganization gave opportunity for creating play groups or social groups according to each child's individual preference. Age-differences in the amount of fiction play stimulated by the picture were striking. Twenty-five per cent of the responses of the two-year-old children were characterized by fiction play, 61 per cent of the responses at three years and 97 per cent of the responses at four years.

The relative constancy in amount of fiction play characteristic of the three age groups in response to the three pictures designed primarily for expressing creative control of human relationships, Pictures 2, 3, and 4, may be seen in Table 8.

TABLE 8
PERCENTAGE OF RESPONSES CONTAINING FICTION PLAY

	2 years (No. = 20)	3 years (No. = 28)	4 years (No. = 29)
Picture 2 (living room)	20	64	93
Picture 3 (bedroom)	35	79	93
Picture 4 (playground)	25	61	97

4. *Personal-Social Interests*

The reorganizations of Picture 4 (Playground) gave opportunity for a creative manipulation of social relationships. Among the three- and four-year-olds, 45 rearrangements were obtained and were classified as follows: (*a*) Separate play groups for girls and for boys (29%); (*b*) a tentative experimentation with mixed play groups as, the boys in the center of a group, and a girl on each side, or the girl in the center of a group, with a boy on each side (27%); (*c*) mixed play groups, each containing a boy and a girl (31%); (*d*) unclassified (14%). The preferred pattern is a social grouping which does not entirely separate the play of the girls from that of the boys. Further insight into children's personal-social interests may be gained by careful observation for each age group of the number of times an absent person or situation is sufficiently interest-compelling to be projected into the interpretation of a picture.

In tabulating such instances, responses have not been included if some characteristic of the absent person or thing is actually present in the picture stimulus, as in the following examples:

Barbara, Picture 3:

"A teddy-bear, a black one, not a green one. I have a pandie."

Keir. Picture 8:

"He's got a teddy-bear. Mine's a yellow one, so big."

Donald. Picture 5:

"My room! I have my room at home."

The reports of the 20 two-year-old children contain 11 references to absent people or things. The following examples are typical for this group:

Delight. Picture 4:

"A ball, no mommie, a tree, two trees."

Jimmie. Picture 10:

"Dat's a little girl. Dat's sleeping. Little girl's mama's sick."

John. Picture 11:

"Little girl, a ball and there's a ball, a ribbon! Her mama gone to store."

Of the 11 references to absent people or things, six concern the mother. There are two references to *"baby,"* one each to *"Daddy," "dollie,"* and *"circus."*

The reports of the 28 three-year-old children contain 22 references to absent people or things. Four references concern the mother. There are two references to *"Daddy,"* five to *"baby,"* two to contemporaries, two to animals, two to the picture of the birthday party, and five to miscellaneous objects.

Typical reports for the three-year-old group are the following:

Gale. Picture 5:

"I'll give the book to my favorite boy. He's downstairs. I play with him." (*Comment on rearrangement of the picture.*)

Carlo. Picture 10:

"He's lying down. I wonder if he mother's dead. Dat moon."

Marilyn. Picture 15:

"That's blue. She's going to a party. That's white. That's brown. Now another! I want to look at the baby way up there!"

The reports of the 29 four-year-old children contain 24 references to absent people or things, five of which are directed to the mother. There are four references to the father, four to contemporaries, three to *"baby,"* two to landscapes, and one reference each to six different objects or situations.

Characteristic references for the four-year-old group are the following:

Helen. Picture 4:

> "They have no birthday cake and no one to make it."

Teddy. Picture 9:

> "A boy and her mother eating dinner. Has she a Daddy? She's got a Daddy, and her Daddy got work."

Jonathon. Picture 14:

> "Oh, a little boy sitting on a bench and a little girl's standing up with a doll. This is some grass and behind is the mountains and behind it some bay."

A survey of exclusions from the *Birthday Party*, Picture 1, in answer to the question, *"Who will come to the party,"* is enlightening from the point of view of its contribution to our understanding of the personal-social development of the individual child, and will be referred to in the discussion of individual protocols.

In addition to such major tendencies in child behavior as identifications and the more generalized social and social-sexual interests of early childhood, other perhaps less strongly motivating interests, such as specific wishes, toy preferences, and color appreciation are all projected clearly into the children's perceptual responses to individual pictures.

Specific interests of nursery school children were brought out clearly in the responses of all three groups to the last question asked in connection with Picture 1, the *Birthday Party, "What will (s)he find in this box?"* (square box), *"What will (s)he find in this box?"* (long box). Among the two-year-olds, about half made verbal suggestions of birthday gifts or birthday wishes, the principal gift mentioned being something to eat, as "candy," "birthday cake," "keam," "cookies," "dinner." About half confined themselves to an interested or excited manipulation of the two boxes with attempts to untie the ribbon. Among the three-year-olds, the majority mentioned a gift, and this was less commonly something to eat than a toy such as a doll, play-house, book, fire engine, train, boat, etc. The four-year-old group made prolific suggestions as to gifts; of 50 suggestions, six referred to things to eat, 16 to toys, and 25 to useful or practical things, such as a brush, a pair of garters, a pair of shows a [illegible]. It is an interesting commentary on our culture that children so early in life are conditioned to the utilitarian. Only one three-year-old child and no two-year-old child expressed an interest in the type of gift which is apparently a dominant preference for the four-year-old group, both for the Institute children and for the children in the emergency nursery schools.

Table 9 briefly summarizes toy preferences for the three groups

TABLE 9

PERCENTAGE OF CASES CHOOSING EACH TOY (PICTURE 5)

Toys	2 years (No. = 9) Girls	2 years (No. = 9) Boys	3 years (No. = 12) Girls	3 years (No. = 13) Boys	4 years (No. = 14) Girls	4 years (No. = 13) Boys
Cart	33	[illegible]	25	63	43	69
Doll	33		42		21	—
Blocks	22	[illegible]	25	31	21	8
Book	11		8	8	14	15
Ball				8	—	8

and is based upon the data obtained from the children's responses to Picture 5, Question 2: "*What does (s)he want from the shelves?*" The children's comments as well as their manipulation of toys indicate that the selection of the first toy from the shelves is influenced strongly by the child's own play preferences. The summary is then simply a record of the children's first choice of toy from the playroom shelves. Toy preferences, as estimated "projectively," indicate the cart or wagon as strongest in interest appeal for nursery school boys in each age group. Toy preferences are more varied for girls, the cart and the doll holding equal interest for two-year-old girls, the doll holding most interest for three-year-old girls, and the cart or wagon for four-year-old girls.

E. COLOR INTEREST

Analysis of the data reveals evidence of developmental differences in response to colors. Table 10 gives a quantitative record of the responses of the three age groups to the 10 colored pictures in the experimental series. In an effort to report the data objectively, only specific references to color, as color-names or the definite use of the word color are counted.

This summary suggests a gradual increase in color interest for each age group, although this may of course depend in part upon an increase in vocabulary.

TABLE 10
COLOR REFERENCES

	No. of color references	Per cent of group responding specifically to color
(No. = 20) 2 years	13	35
(No. = 28) 3 years	18	36
(No. = 29) 4 years	26	48

F. PROJECTION

While the present report is not concerned directly with the discussion of the significance of the protocols as projective data, examples already given will serve to indicate some of the possibilities in this field. Attention may be called to the reports obtained from Clifford, summarized above. Clifford's responses strongly suggest a preoccupation with "bad" behavior; the nature of this preoccupation, and its implications concerning relationships with parents or other persons in his environment, can of course be ascertained only through other clinical data. A rather obvious example of a simple type of projective tendency is found in the following protocol from Michael:

> Michael, 2 years, 9 months, who had cried at the prospect of coming indoors, but was now enjoying the pictures.
>
> *Picture 9: (Child Refusing to Eat)*
>
> "Here's anudder boy got all fru crying. Dat's anudder mommie. He won't go for a walk with his mommie washing dishes. Are the dishes all done?"

G. SUMMARY

The analysis of the data from the point of view of age-trends indicates:

1. Developmental changes in the perceptual reactions of children to pictures.

a. An increase with age of the tendency to interpret picture content in terms of activity, rather than in terms of static form, and in terms of inner as well as outer activity.

b. A sequence of development from literal, concrete interpretation of details to more subjective interpretation.

c. A sequence of development in part-whole balance, characterized by the following stages:

(1). Interpretation of a part as the whole.
(2). Unanalyzed whole.
(3). Incomplete whole incompletely analyzed.
(4). Complete whole incompletely analyzed.
(5). Complete whole completely analyzed.

2. Developmental changes in motivation and interest.

a. Progressive changes in the identification process.
b. Broadening of social interests.
c. Progressive structuring of interests, as represented in wishes, toy preferences, and color responses.

Against this background of typical behavior patterns, individual perceptual reactions can be more readily understood.

IV. INDIVIDUAL PROTOCOLS

It is the purpose of this section of the report to illustrate, through a selection of individual records, the nature of the data obtained in particular cases and to discuss the contribution of the data to an understanding of the inner world of the individual child.

The following record is that of Bertha, a little girl of 3 years and 6 months, *IQ* 92. Three experimental periods instead of the usual two were necessary for obtaining the data.

1. *Birthday Table*

Examiner: Let's pretend it's someone's birthday; whose birthday is it?

"This one." (Selects girl doll with pink dress, after first picking up a girl doll in one hand, and a boy doll in the other.)

Examiner: Who will come to the party?

(Picks up the lady, then the man.) *"Is this a man?"* (Pointing to doll representing a father.)

What will (s)he find in this box? (square box).

"A candle!"

What will (s)he find in this box? (long box).

"A candle!!" excitedly.

2. *Living-Room, Opening Out-of-Doors*

1. What is this picture about? *"I want the birthday party!"* (Picks up the dog. Looks at picture very closely.) *"Are your teeth coming up?"*

2. Can you make another picture? Becomes interested in sheen of the cellophane covering, and looks for own image, sticking out tongue.

3. *Bedroom, Opening Into Smaller Room*

1. What is this picture about? *"What's this,"* pointing to baby. *"What's this."* (Picking up ball.) *"Can I cut it? Can I throw it?"*, after being told it is a ball.

2. Can you make another picture? *"Have you a tail? Is this Mommie?"*, touching experimenter's sleeve several times.

4. *Playground*

1. What is this picture about?

"Where's the birthday?"

2. Can you make another picture?

"A swing." Sweeps all the movable pieces off the background, with comment, *"Somebody took my wagon!"*

5. *Playroom, with Toys*

Whose playroom shall we call this? (No response.)

What does (s)he want from the shelves?

Picks up all the toys, beginning with blocks, with question, *"What dat?"*

Can you make another picture with this child and these playthings?

Leans close down over picture, apparently enjoying the reflection.

Sticks out tongue against cellophane. Turns picture on back side, with comment, *"Can't see."*

Pictures in Silhouette and in Color

What is this picture about? What does it make you think of? Let's tell a story about the picture.

6. *Sulky Child*

"Lady!"

7. *Mother Coercing Child*

"Dat's a lady. Dat's a little boy!"

8. *Mother Caring for Baby in Crib*

"Bed and is there a mattress? Do they have mattresses under beds? And do they have tails?"

9. *Child Refusing to Eat*

Looks closely. "What is dat? Those are boys and a girl."

10. *Child Sleeping*

"She's sleeping there in bed. That's just like my bed. Do you have a tail?"

11. *Balloon Play*

"He has socks just like mine and black sandals."

12. *Two Children with Dog*

"What is dis?" Looks long and quietly. "What is dat?" Want to go out! That isn't nuffin."

13. *Father Tossing Baby*

"That is red trousers on."

14. *Daydreaming*

"That is a seat," pointing to green. "A girl," pointing to doll. "A lady there," pointing to girl.

15. *Big boy and little girl*

Looks closely. "That is a girl and a boy."

A comparison of this child's perceptual responses with those of other children of nursery school age, indicates that she is at a *relatively low level of apperceptive maturity.* A striking characteristic of the responses to both silhouettes and the colored series with simple emotional appeal, is the static quality of interpretation. Activity is suggested in the description of only one of the 10 pictures, Picture 10: "*She's sleeping there in bed.*"

It will be recalled that among the three-year-old children studied, 37 per cent of the responses to Pictures 6-15 involved static form. Among two-year-olds the corresponding percentage was 73. In Bertha's case, 90 per cent of the responses are in terms of static form. Moreover, the interpretations are invariably literal and concrete. There is not a single instance in which the overt, outward fact is given an inward or symbolic meaning.

A study of the perceptual responses from the point of view of part-whole balance indicates that in seven pictures, the two lowest stages of perceptual grasp are represented. Perceptual reactions to Picture 8 and to Picture 13 are organized about a detail which is substituted for the whole, as:

Picture 8:

> "Bed and is there a mattress? Do they have matresses under beds? And do they have tails?"

Picture 13:

> "That is red trousers on."

The response to Picture 12 shows no organization at all:

> "What is dis? What is dat? Want to go out! That isn't nuffin."

Responses to Pictures 6, 7, 9, and 15 represent the unanalyzed whole type of perception, with an error in the whole reported for Picture 9.

Pictures 11 and 14, each containing only two central figures, are perceived as incomplete wholes, incompletely analyzed.

Picture 10, containing a single central figure, is perceived as a whole, but is incompletely analyzed. The three major details, flying curtains, moon or sun shining in the window, and the picture on the wall are all omitted from the description.

These three characteristics of the perceptual reactions to Pictures 6-15, the dominance of static form interpretation, the literal concrete emphasis, and the incompleteness and inaccuracy of perceptual patterns, suggest an inner experience which is conspicuously uncreative, undeveloped, and unstructured.

The amount of fiction play, and the evidence of social interest are also closer to the two-year than the three-year level. In the response to Picture 4, the sudden comment, *"Somebody took my wagon,"* followed by a sweeping movement which pushed all the movable figures from the picture, gave an excellent example of the child's aggressive reaction to frustration.

In Bertha's record certain additional typical characteristics stand out sharply; the turning of the child's attention to the cellophane and the self-image reflected there; the persistence, over a time interval of four weeks, of such questions as: *"Are your teeth coming up?"*; *"Have you a tail?"*; *"And do they have tails?"*; *"Do you have a tail?"*; the uncertainty prompting the question, *"Is this Mommie?"* asked of the experimenter.

These highly individualistic characteristics of the report suggest a striking preoccupation of thought and feelings with personal, inner stimuli. The distraction of attention from present objective stimuli, the quality of the perceptual or intellectual experience and the underdevelopment of social feeling create an impression of extreme deviation from the norm.

The second record selected for consideration is that of a boy, Rudolph, 4 years and 7 months of age, *IQ* 121.

1. *Birthday Table*

Let's pretend it's someone's birthday. "Mine is September 13th. I'm four and a half."

Whose birthday is it? Child selects girl doll with yellow dress.

Who will come to the party? Child brings in blue boy. "That's a pretty boy, isn't it? This man. This one with the pink dress. This one with green pants. This too" (the mother doll).

What will (s)he find in this box? (Square box). "I think she'll find a toy."

What will (s)he find in this box? (long box). "And I think that will be a brush. There are one, two, three, four candles and I'm four and a half."

2. *Living-Room (opening out-of-doors)*

What is this picture about?

"A dog! Oh, there's the girl, and there's the boy again.

Oh yes, she's four years old. I think the dog eats bones, and the girl and boy eat dinner."

Can you make another picture?

Places boy in chair, dog outside, and girl standing.

3. *Bedroom (Opening Into Smaller Room)*

"Is that a very pretty picture," asked before picture was shown.

What is this picture about?

1. "I think this is about a boy (baby) throwing a ball to this pretty boy!"

Can you make another picture?

2. "The boy's laying down and the daddy's laying down. The girl is standing by her mommie."

4. *Playground*

What is this picture about?

1. Little trees and a house. I think I'll climb up it! I think they're playing football. You have to kick with this foot. I can kick it pretty high. I can kick with this foot, too. Oh, I like this boy best! He's pretty. So's the ball."

Can you make another picture?

2. "They're marching (blue boy, yellow girl, pink girl, green boy). She's got a pocket. They're marching this way," noticing that profile of boy in green faces a direction opposite to that of the other three children.

5. *Playroom with Toys*

Whose playroom shall we call this? Selects boy in blue.

What does (s)he want from the shelves?

"I think he wants the wagon."

Can you make another picture with this child and these playthings?" "I think I'll put this block in the wagon." Puts all the blocks in the wagon. "Pull, pull, pull!"

Pictures in Silhouette and in Color

What is this picture about? What does it make you think of? Let's tell a story about the picture.

7. *Mother Coercing Child*

"It looks like a woman. No, it looks like Snow White."

7. *Mother Coercing Child*

"Looks like a woman and a little girl. Taking her to the hospital. I was born in the hospital in Oakland."

8. *Mother Caring for Baby in Crib*

"The woman's in the bed and the baby's in the bed. This girl is going to collapse. My daddy collapsed."

9. *Child Refusing to Eat*

"This is about a woman feeding the boy."

10. *Child Sleeping*

"Oh, I think this is about a boy sleeping, and the moon is coming in and the window is open, and the curtains are waving. The boy is sound asleep. I like sleeping. I sleep well."

11. *Balloon Play*

"That's a new picture. That's purple! That's a girl holding a balloon and a boy stepping on it. Does he want to break it?"

12. *Two Children with Dog*

"Oh, this boy's pretty! The boy is walking on the sidewalk: a girl and a dog."

13. *Father Tossing Baby*

"Look at that man put the baby up. A little, little child with his clothes off," pointing to older child. "They're walking on the street and they'll get run over."

14. *Daydreaming*

"Oh yes! boy sitting on there and a girl holding a doll. Flowers!"

15. *Big Boy and Little Girl*

"Oh, that's a pretty picture! Boy telling girl, 'You'd better go across the street!' She got run over too. Anything more you got for me to do?"

In contrast to the first record discussed, this child interprets all of the pictures from Picture 7 through Picture 15 in terms of activity. Picture 6, *Sulky Child*, is the only picture in the series of 10 calling for verbal description, for which an interpretation in terms of static form is given. In at least four of the pictures, the activity projected is an inner activity, rather than an overt activity, as the suggestion of purpose in Picture 7, *"taking her to the hospital"*; the suggestion of inner state in the case of the little girl in Picture 8, *"this girl is going to collapse"*; the reference to purpose or mood in the case of the little boy in Picture 11, *"does he want to break it?"*; and the thought of the little boy in Picture 15, expressed in the words, *"You'd better go across the street."*

A comparison of this four-year-old boy's pattern of interpretation for the 10 pictures for which group averages were computed, with the average pattern for the 29 four-year-old boys studied, gives evidence of superiority in performance, from the point of view of creative mental activity.

Rudolph's interpretations show less static form (10 per cent *vs.* 23 per cent) than was found for the average of 29 boys of his age group (Pictures 6-15). Forty per cent of his interpretations are in terms of inner activity, as compared with 25 per cent in the group average of four-year-olds.

The record shows also a good degree of mental maturity as measured by progress from an entirely concrete or literal interpretation of the objective aspects of the stimulus to an understanding of objective details as symbols in a more meaningful whole. For example, the child's posture in Picture 7 is perceived as symbolic of an inward state of reluctance or unwillingness and hence the projection, by the child under discussion, of the hospital destiny for the child coerced. Similarly the objective details of Picture 10, the moonlight, the open window, the moving curtains as well as the sleeping figure are all interpreted as symbols of rest and refreshment, and so the expression of satisfaction in the comment, *"I like sleeping. I sleep well."* The little boy's posture in the Balloon Picture is understood as symbolic of aggressive tendency, and therefore the question, *"Does he want to break it?"*

When the report is considered from the point of view of part-whole balance in perception, no responses are found which belong to the three lower categories of perceiving, the perception of a detail in place of a whole, an unanalyzed whole, or an incomplete whole incompletely analyzed. Both pictures containing three central figures are described with some detail. The descriptions of all 10 pictures must be classified as complete wholes analyzed either incompletely or completely with reference to major details.

Again, the pictures with movable parts contribute the most satisfactory data for an interpretation of motivation and of personal-social development. The direction of identifications cannot be definitely stated on the basis of this record. Whereas, the little boy in blue is given the playroom and toys and the ambiguous figure in Pictures 9 and 10 is interpreted as a boy, a little girl is made the heroine of the birthday party, and the child figure in Pictures 7 and 8 is described as a little girl. In Picture 13, the equivocal figure is referred to as *"a little, little child."* Interest is therefore not narrowly self-limited.

The response to Picture 2 shows a real interest in the children as well as in the dog, and pleasure in finding them again: *"Oh there's the girl and there's the boy again. Oh yes, she's four years old."* The reorganization of the picture illustrates the somewhat typical reaction of the boys, that of giving the comfortable big chair to the boy doll, placing the dog out of doors and letting the little girl stand up.

The reaction to Picture 3, as well as to Picture 2, may indicate a readier identification on the child's part with the boy doll and perhaps with the father doll as well, inasmuch as in the rearrangement of this picture the little bed is given to the boy doll and the big bed to the daddy doll, with the comment, *"The boy's laying down and the daddy's laying down. The girl is standing by her mommie."*

Two reorganizations were spontaneously given for the Playground Picture. The first reorganization suggests the mixed play group of boy and girl, and the second arrangement, the intermediate pattern, four children marching in a row, the girls inside, the boys outside.

The fiction play, for all three pictures, Nos. 2, 3 and 4, indicates a high level of social interest. A broad social interest as well is suggested by the use, in reorganizing the picture, of the two more mature play patterns, the tentative mixed play pattern, as described above, and the boy-girl pattern.

Specific interests stand out clearly in the report. The interest in toys, but also in useful things, is shown in the choice of the birthday gifts. The boy's interest in the wagon is projected into the playroom scene: *"I think he wants the wagon."*

A tendency to aesthetic judgment and appreciation of color are outstanding characteristics of the report. The word "pretty" occurs half a dozen times.

Picture 1:

"That's a pretty boy, isn't it?"

Picture 3:

"Is that a very pretty picture? I think this is about a boy (baby) throwing a ball to this pretty boy!"

Picture 4:

"Oh, I like this boy best! He's pretty. So's the ball."

> *Picture 12:*
>
> "Oh, this boy's pretty!"
>
> *Picture 15:*
>
> "Oh, that's *a pretty picture!*"

The flowers, observed in Picture 14, were mentioned in only four out of 77 descriptions of this picture. Colors are specifically mentioned three times in the report. Only two other boys out of the 39 boys studied have as high a score for color interest.

In a report indicating superiority in mental capacity, well developed social interest, and unusual sensitivity to beauty and to color, it is surprising to find the recurrence of the following theme.

> *Picture 7:*
>
> "Looks like a woman and a little girl. *Taking her to the hospital.* I was born in the hospital in Oakland."
>
> *Picture 8:*
>
> "*The woman's in the bed and the baby's in the bed. This girl is going to collapse.* My daddy collapsed."
>
> *Picture 13:*
>
> "Look at that man put the baby up! A little little child with his clothes off. They're walking on the street and *they'll get run over.*"
>
> *Picture 15:*
>
> "*O that's a pretty picture! Boy telling girl,* 'You'd better go across the street!' *She got run over too.*"

The report brings out clearly a strain of anxiety which is already affecting the child's adjustment to the nursery school environment, where he is regarded as a poorly adjusted child. A partial explanation is probably to be found in home conditions. The child's mother has been threatened for many months with a nervous breakdown. The child has recently had a tonsillectomy and while at the hospital was visited by his father, who fainted upon entering the hospital room. The mother frequently threatens the child when naughty with a return to the hospital and the father's collapse.

A third record which may help in evaluating a projective technique, based upon pictures, as a method of understanding child personality, is the record of another boy of the same age and approximately the same *IQ* as Rudolph. Peter is 4 years and 7 months of age, with an *IQ* of 125.

1. *Birthday Table*

Let's pretend it's someone's birthday.

Whose birthday is it? Peter selects the boy in blue.

Who will come to the party? "These two and the little girl. Mother and Daddy."

What will (s)he find in this box? (*square box*). "Easter egg."

What will (s)he find in this box? (*long box*). "A shovel."

2. *Living-Room Opening Out-of-Doors*

What is this picture about?

1. "A little boy and girl with a dog and the window open."

Can you make another picture?

2. "I think they would like to be outdoors. She has brown hair."

3. *Bedroom (Opening Into Smaller Room)*

What is this picture about?

1. "A mother and a daddy, and a little baby and a little girl and a little boy and a pond! A star fish and a rooster."

Can you make another picture?

2. "I want a picture with just a boy and a girl and a father, not any mother, cause I want to pretend the mother has gone to work and the father stays home. Pretend the mother has gone and lives somewhere else. Maybe they will be lonesome. Only two people now. The little girl has gone with the mother. Now the father and the little boy has gone to find them. They forgot their things, didn't they?"

4. *Playground*

What is this picture about?

1. "There, they're all by themselves outdoors. She doesn't play ball," referring to doll in pink dress.

Can you make another picture?

2. "Pretend someone kicks it up, up in the sky. It might go over in the field," pointing to the trees.

5. *Playroom, with Toys*

Whose playroom shall we call this? "This little boy's because it was his birthday."

What does (s)he want from the shelves? "I think he would like the best this cart." Removes everything from the shelves.

Can you make another picture with this child and these playthings? Careful distribution of toys among the dolls.

Pictures in Silhouette and in Color

What is this picture about? What does it make you think of? Let's tell a story about the picture.

6. *Sulky Child*

"A little girl dancing."

7. *Mother Coercing Child*

"They're going for a walk. They're going through a tunnel."

8. *Mother Caring for Baby in Crib*

"Mother's fixing a little boy, and the little boy is watching. He's got a teddy-bear. You're writing things I told you."

9. *Child Refusing to Eat*

"I see a little girl. Mama makes the little girl eat her dinner, but she doesn't."

10. *Child Sleeping*

"The little boy's asleep and a storm blowing the curtains and old Mr. Moon is looking at the storm."

11. *Balloon Play*

"Dey little boy trying to pop his balloon and the little girl isn't popping hers. She's just holding hers. He's popping it. He doesn't want it."

12. *Two Children with Dog*

"They're running somewhere and the little dog's running after. The little boy's crying because the little girl's running faster, running down the street."

13. *Father Tossing Baby*

"He's putting the baby up in the air. The little boy's holding the Daddy's pants."

14. *Daydreaming*

"The little girl's holding the doll, and the little boy's sitting on a tree that's broke down. There's green grass and another tree."

15. *Big Boy and Little Girl*

"A little boy came in and found a little girl helping her mother cook."

This record, like the preceding record, indicates a high level of mental activity. All 10 of the pictures presented for verbal description are interpreted in terms of activity, and at least three in terms of inner activity, as Picture 9 in which reference is made to the child's resistance or obstinacy, *"Mama makes the little girl eat her dinner, but she doesn't"*; Picture 11, in which a motive for popping the balloon is suggested, *"He's popping it. He doesn't want it,"* and Picture 12 in which a reason is offered for the little boy's crying: *"The little boy's crying because the little girl's running faster—running down the street."* It is difficult to determine whether the interpretation for Picture 15 suggests simple overt activity, or carries the meaning of the little girl's sympathetic sharing of the mother's activity. In all doubtful cases, the interpretation has been classed as one of overt activity.

The pattern of interpretation for this record may therefore, be expressed as (*a*) no static form, (*b*) 70 per cent outer activity; (*c*) 30 per cent inner activity; in contrast to the average for the age group, which is 23 per cent static form, 51 per cent outer activity, and 25 per cent inner activity. In at least three of 10 pictures, the report makes a symbolic interpretation of objective details. In Picture 9, the posture of the child is interpreted in terms of negativism and resistance. The objective signs or symbols of destructive aggression and of distress are recognized as such in Pictures 11 and 12.

An analysis of the responses from the point of view of part-whole balance indicates that perception, under the conditions of the experiment, is organized in every instance as a complete whole, either completely or incompletely analyzed.

It is clearly evident from a consideration of the three aspects of the report already discussed that the record shows creative, alert intelligence.

An analysis of the data obtained from the pictures with movable parts, is enlightening in this instance, especially for its contribution to an understanding of motivational tendencies and personal-social development.

In five of the seven experimental situations in which an identification of some sort is the natural response, the report shows a tendency toward more sympathetic interest in *"Daddy"* and in little boys than in mother or little girls. A boy is selected as central figure for the birthday party. The choice of central figure for the playroom is also, *"This little boy, because it was his birthday."* The ambiguous figure in Picture 8 is perceived as a little boy: *"The little boy is watching; he's got a teddy-bear."* In Picture 10, the sleeping figure is called a little boy, and the older child in Picture 13 is also described as a little boy. *"The little boy's holding the Daddy's pants."* The exception is found in the interpretation of Picture 9, *"I see a little girl. Mama makes the little girl eat her dinner, but she doesn't."* The response to Picture 7 is equivocal: *"They're going for a walk. They're going through a tunnel."*

The fiction play in the reorganization of Picture 3 brings out very clearly the child's emotional identification with his father. *"I want a picture with just a boy and a girl and a father, not any mother, cause I want to pretend the mother has gone to work and the father stays home. Pretend the mother has gone and lives somewhere else. Maybe they will be lonesome. Only two people now. The little girl has gone with the mother. Now the father and the little boy has gone to find them. They forgot their things, didn't they?"*

The objective situation underlying this phantasy play is, for the child concerned, a broken home experience. Father and mother are divorced. The child and his little sister live with the mother, who works outside the home. The child's longing for the father, the fear of loneliness without the mother—*"Maybe they will be lonesome"*—the experiencing in imagination of companionship with Daddy alone—*"The little girl has gone with the mother,"* and the final wish-fulfillment of a reunited family—*"Now the father and the little boy has gone to find them,"* are all vivid indications of an emotional tension which has not yet been resolved in this child's inner world.

The fantasy play which suggests the affection of the little girl for the mother perhaps explains the interpretation of Picture 9 as a mother making her little girl eat dinner, and the interpretation of Picture 15, *"A little boy came in and found a little girl helping her mother cook."*

In response to Picture 4, the comment is made, apparently with some relief from the conflict set up by the bedroom scene, *"There, they're all by themselves outdoors."* The child's rearrangement of the picture shows two groups, in opposite corners of the playground, each group containing a little girl and a little boy.

The three complete records presented above, in addition to the excerpts which are quoted throughout the report, illustrate the type of data which may be expected from the use of pictures in a projective study of nursery school children. The normative and interpretive discussion of these protocols has been for the most part "peripheral" rather than "central" in emphasis. That is to say, the responses to the pictures have been examined from a relatively "exterior" point of view, with little attempt to diagnose the deeper aspects of personality structure. The writer feels that if such analyses are desirable they should be deferred until wider experience has been obtained in the use of this procedure.

V. PRELIMINARY EVALUATION OF THE PROCEDURE

A projective technique based upon picture materials seems particularly appropriate for children of nursery school age for three reasons. The first is the relative lack of self-consciousness in children of this age. Only one child in the entire group of 77 children seemed interested in the recording of responses, and commented as follows: (Peter. Picture 8): "*You're writing things I told you.*" In the second place, the absence of fixed habits of perceiving in little children increases the possibility of some degree of self-projection in the organization of a perception. Finally fantasy itself, which constitutes the scientific data of projection studies is, in the form of fiction play, one of the most characteristic and universal activities of children at three years and four years of age. It is apparent that because of verbal difficulties or for other reasons, the pictures are not well adapted for eliciting fantasy among two-year-olds.

The pictures with movable parts were the most thoroughly enjoyed by the children. Provision for manipulation of the pictures, as well as for verbal comment, proved a good method of introducing the child both to the experiment and to the series of pictures. Because of their content, and the invitation to an active type of response, they proved valuable in revealing personal characteristics not as readily shown in procedures requiring merely a verbal reaction.

The silhouettes presented the most difficult and the most unfamiliar type of picture material. In the total group there were 12 refusals to interpret the silhouettes, as compared with two refusals to interpret the colored pictures presented for verbal description. The following comments are typical of the children's reactions to the series of silhouettes:

Darleen. Picture 6:

"I don't like these funny pictures. This is a boat up here" (pointing to moon in little picture). "This is a girl, but I don't *know what's putting her leg up for.* She might be dancing."

Gail. Picture 6:

"I won't tell you. It's not a very good picture."

Sandy. Picture 6:

"A girl. It's so black, I can't see it."

Jimmie. Picture 7:

> "We can see Mama. No, we can't see Mama," pointing to face.

Though often confusing to the children, the silhouettes may elicit valuable projective material, through the element of ambiguity thus introduced.

The colored pictures, presented after the silhouettes, were welcomed by the children. These were the pictures most often described as "pretty."

It is, however, difficult to evaluate the three types of pictures separately. In general they have derived their value as stimuli largely from the fact that they were carefully planned to approximate the life experiences of little children, to interest children and so to become the medium of projection for "the meanings, significances, patterns, and feelings" (3, 402) which constitute the beauty and the mystery of human personality.

VI. CONCLUSIONS

The analysis of the data obtained through this projection experiment from the point of view of age-trends, brings out very clearly certain definite developmental tendencies in perceptual reaction to pictures, which have important implications for an understanding of both intellectual and motivational growth throughout the preschool period.

Such developmental tendencies in perceptual reaction to pictures are:

1. The tendency to interpret the visual stimulus first in terms of static form, and later in terms of activity. Interpretation in terms of inner activity, thought or feeling, is a more mature pattern of interpretation than interpretation in terms of outward activity.

2. The tendency to interpret a visual stimulus at first concretely or literally, with slow development toward more subjective interpretation.

3. The tendency to interpret a visual stimulus first in terms of a primitive unanalyzed whole, or a detail standing for the whole, with gradual development of the capacity to enlarge the whole and to increase the detail. The second and third tendencies are difficult to measure quantitatively because of the qualitative differences involved.

4. The tendency toward some degree of self-identification with picture stimuli, or the projective tendency itself, which is most clearly though indirectly apparent in the fiction play of children.

Against this background of general developmental tendency, individual differences in reaction, which are of the very essence of personality, stand out sharply.

The technique developed in this experimental study with pictures has yielded not only descriptive data useful for understanding age trends, but also something of interest and of unique significance in the perceptual reactions of each one of the subjects studied. So long as an "individual's private world of meanings, significances, patterns and feelings" (3, P402) is valued, it appears worth while to continue investigation of this difficult field by every method which suggests a possibility of meaningful results. The method presented in this preliminary report is sufficiently promising to justify further use, with the purpose of establishing principles of interpretation that may apply to deeper aspects of personality as well as to the more superficial aspects considered here.

REFERENCES

1. Baruch, Dorothy W. Play techniques in preschool as an aid in guidance. *Psychol. Bull.*, 1939, **36**, 570.
2. Boring, E. G., Langfeld, H., & Weld, H. P. Introduction to Psychology. New York: Wiley, 1939.
3. Frank, L. K. Projective methods for the study of personality. *J. of Psychol.*, 1939, **8**, 389-413.
4. Horowitz, R., & Murphy, L. B. Projective methods in the psychological study of children. *J. Exper. Educ.*, 1938, **7**, 133-140.
5. Klopfer, B. The technique of the Rorschach performance. *Rorsch. Res. Exch.*, **2**, 1-14.
6. Murchison, C. Handbook of Child Psychology. Worcester, Mass.: Clark Univ. Press, 1933. Pp. 956.
7. Murphy, L. B. Social Behavior and Child Personality. New York: Columbia Univ. Press, 1937.
8. Murray, H. A. Explorations in Personality. New York: Oxford Univ. Press, 1938.
9. Stern, W. Cloud pictures; a new method for testing imagination. *Charac. & Personal.*, 1937, **6**, 132-146.
10. Symonds, P. Criteria for the selection of pictures for the investigation of adolescent fantasies. *J. Abn. & Soc. Psychol.*, 1939, **34**, 271-274.